

**CLONING AND CHARACTERISATION OF *DROSOPHILA*
IMP DEHYDROGENASE**

ROGER B. SLEE

**Thesis presented for the degree of Doctor of Philosophy
University of Edinburgh
1992**



ABSTRACT

A preliminary molecular characterisation was conducted of a transcription unit located adjacent to an enhancer-trap transposon p[lac Z, ry⁺] at 38C, in a transformant exhibiting testis-specific expression of the β -galactosidase reporter gene.

Northern analysis with a probe (A-5.1) consisting of a genomic fragment situated immediately upstream of the transposon, identified a region which encoded three transcripts with distinct sizes and expression profiles, including 1.0 kb and 2.5 kb species which are highly enriched in adult ovary RNA. The same probe was used to isolate a cDNA clone (OV.1) representing one of the transcripts detected by A-5.1. In Northern hybridisation OV.1 detected all the same transcripts as A-5.1 plus an additional testis-specific transcript of 0.8 kb. An OV.1 probe was used to isolate a cDNA (T1) corresponding to the testis-specific transcript and T1 was shown to recognise all the same transcripts as OV.1 on a Northern blot except for the 1.0 kb ovary enriched species. OV.1 and T1 were sequenced and found to represent overlapping transcripts from the same gene. OV.1 contains unique 5' sequences which are encoded immediately upstream of p[lac Z, ry⁺] whilst the region common to OV.1 and T1 is encoded more than 8 kb distant in the genome. The picture that emerged from this analysis was of a gene which included an intron of over 8 kb and generated multiple, independently regulated, overlapping transcripts including testis-specific and ovary enriched variants.

The predicted protein sequences of T1 and the overlapping region of OV.1, were found to be highly similar to the C-terminal portion of inosine 5' monophosphate dehydrogenase (IMPD), an enzyme involved in *de novo* purine nucleotide synthesis, from a range of species. A third cDNA clone (T3) which represents the 2.5 kb transcript and overlaps with OV.1 and T1 was isolated and sequenced and found to encode the complete *Drosophila* IMPD.

The conclusion from this research was that the *Drosophila* IMPD gene has a far more complex expression pattern than was predictable on the basis of the enzyme's known functions and molecular data from other species. The possible significance of the observed expression pattern is considered.

ACKNOWLEDGEMENTS

I would like to acknowledge everyone who has helped and supported me throughout my research project.

My supervisor, Mary Bownes and all members of the Bownes group, especially Alberto, Nian and Brenda.

Frank Johnston and Graham Brown for their excellent photographic work.

My mother and father, Sibylle and Jocky.

ABBREVIATIONS

amp	Ampicillin
ATP	Adenosine-5'-triphosphate
bp	Base pair
°C	Degrees centigrade
cdNA	Complementary deoxyribonucleic acid
Ci	Curies
cm	Centimetre(s)
cpm	Counts per minute
(d)dATP	2'(3'-di) Deoxyadenosine-5'-triphosphate
(d)dCTP	2'(3'-di) Deoxycytosine-5'-triphosphate
(d)dGTP	2'(3'-di) Deoxyguanosine-5'-triphosphate
dNTP	Deoxynucleotide-5'-triphosphate
(d)dTTP	2'(3'-di) Deoxythymidine-5'-triphosphate
dH ₂ O	Distilled water
DEAE	Diethylaminoethyl
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DTT	Dithiothreitol
EDTA	Diaminoethanetetra-acetic acid
fs	Female sterile
FSB	Formaldehyde sample buffer
g	Gram(s)
[³ H]	β-emitting isotope of hydrogen
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid

kb	Kilobase
Klenow	Large fragment of DNA polymerase I
Krpm	Kilorevolutions per minute
λ	Lambda bacteriophage
L	Litre
M	Molar
mA	Milliampere(s)
mAP	Messenger Affinity Paper
mCi	Millicurie(s)
mg	Milligram(s)
ml	Millilitre(s)
mm	Millimetre(s)
mm/L	Millimoles per litre
mM	Millimolar
mmol	Millimole(s)
min	minute(s)
MOPS	Morpholinopropanesulphonic acid
mRNA	Messenger ribonucleic acid
m wt/ml wt	Molecular weight
ng	Nanogram(s)
nmol	Nanomole(s)
OLB	Oligo labelling buffer
OD	Optical density
[³² P]	β -emitting isotope of phosphorous
pers.comm.	Personal communication
PEG	Polyethylene glycol
%	Percentage
pg	Picogram(s)

pH	$-\log_{10}$ (hydrogen ion concentration)
Poly (A) ⁺ RNA	Polyadenylated ribonucleic acid
psi	Pounds per square inch
RNA	Ribonucleic acid
RNase	Ribonuclease
rRNA	Ribosomal ribonucleic acid
rpm	Revolutions per minute
[³⁵ S]	β -emitting isotope of sulphur
SDS	Sodium dodecyl sulphate
Tris	Tris(hydroxymethyl)-amino-methane
Triton-X-100	Octylphenoxypolyethoxyethanol
TCA	Trichloroacetic acid
U	Unit(s)
UTP	Uridine-5'-triphosphate
UV	Ultraviolet
μ Ci	Microcurie(s)
μ g	Microgram(s)
μ l	Microlitres(s)
μ M	Micromolar
μ mol	Micromole(s)
V	Volts
v/v	Volume per volume
w/v	Weight per volume

CONTENTS

Title	i
Declaration	ii
Abstract	iii
Acknowledgements	v
Abbreviations	vi
Contents	ix

CHAPTER ONE

INTRODUCTION

1.1	Somatic sex-determination	2
1.1.1	The search for genes downstream of the regulatory hierarchy	5
1.1.2	The yolk protein genes	6
1.1.3	Gonadal sex-specific gene expression	6
1.2	Germline sex-determination	7
1.3	Gonadogenesis/gametogenesis	10
1.4	Description of reproductive organs and gametogenesis in both sexes	11
1.4.1	Male	11
1.4.1.1	Spermatogenesis	11
1.4.1.2	Spermogenesis	12
	a) pre-elongation	12
	b) elongation	13
	c) individualisation	13
	d) coiling	14
1.4.2	Female	14
1.4.2.1	Oogenesis	14
1.5	Development of the reproductive organs	17
1.5.1	Gonadogenesis	17

1.5.2	Formation of the adult reproductive system	19
1.6	Mosaic analysis in the study of development	20
1.7	Surgical techniques	22
1.8	Genetic regulation of development of the reproductive system	23
1.9	Genetic and molecular analysis of spermatogenesis	23
1.9.1	The Y chromosome	25
1.9.2	Molecular studies in spermatogenesis	26
1.9.3	The $\beta 2$ tubulin gene	30
1.10	Genetic and and molecular analysis of oogenesis	31
1.10.1	Ovary transplantation	32
1.10.2	Pole cell transplantation and somatic recombination	32
1.11	Ovary non-autonomous mutations	33
1.12	Molecular analyses	35
1.12.1	The yolk protein genes	35
1.12.2	The chorion protein genes	37
1.13	Maternal effect mutations	38
1.13.1	Sex-determination	38
1.13.2	Anterior-posterior axis	39
1.13.2.1	The role of the posterior group genes in pole-plasm formation	43
1.13.3	Dorsal-ventral polarity	44
1.13.4	Determination of terminal pattern elements	48
1.14	Enhancer-trap gene detection system	51
1.15	Cloning of <i>Drosophila</i> IMP dehydrogenase	56

CHAPTER TWO

MATERIALS AND METHODS

2.1	Solutions	58
2.2	Maintenance of <i>Drosophila</i> Stocks and Strains	61

2.3	<i>Drosophila melanogaster</i> Life Cycle - Collection of Developmental Stages	63
2.3.1	Stages of <i>Drosophila melanogaster</i> Development at 25°C	63
2.3.2	Collection of Eggs (0-6 hour embryos)	64
2.3.3	Collection of Late 3rd Instar Larvae and Pupae	64
2.3.4	Collection of Sexed Adults	64
2.3.5	Libraries	67
2.3.5.1	Canton S genomic library	67
2.3.5.2	Testis cDNA library	68
2.3.5.3	Adult male cDNA library	68
2.3.5.4	Ovary cDNA library	68
2.4	Microbial Strains and Media	68
2.4.1	Microbial Strains	68
2.4.2	Media	69
2.5	Methods	70
2.5.1	General Methods	70
2.5.1.1	Phenol Extraction	70
2.5.1.2	Ether Extraction	70
2.5.1.3	Precipitation of Nucleic Acids	70
2.5.1.4	Estimation of DNA and RNA Yields	71
2.5.1.5	Restriction Endonuclease Digestion of DNA	71
2.5.1.6	Extraction of DNA from Agarose Gels	71
2.5.1.7	Ligation of DNA Molecules	72
2.5.1.8	Transformation into <i>E.coli</i>	72
2.5.1.9	Preparation of Dialysis Tubing	73
2.5.2	Preparation of Genomic DNA from Adults of <i>Drosophila melanogaster</i>	73
2.5.3	Preparation of RNA from Different Developmental Stages and Tissues of <i>Drosophila melanogaster</i>	74
2.5.3.1	Large Scale Total RNA Preparations from Eggs, Larvae, Pupae and Adults	74
2.5.3.2	Small Scale Total RNA Preparations from Ovaries, Testes and Carcasses	75

2.5.3.3	Preparation of poly (A) ⁺ RNA	75
2.5.4	Preparation of Plasmid DNA	77
2.5.4.1	Small Scale	77
2.5.4.2	Large Scale	77
2.5.5	Manipulations with Lambda	78
2.5.5.1	Plating of Phages	78
2.5.5.2	Plate Lysates - Phage Storage	78
2.5.5.3	Phage DNA Extraction	79
2.5.5.4	Screening libraries / Plaque Lifts	81
2.5.5.5	<i>In Vivo</i> excision of pBluescript from λ ZAP II	82
2.5.6	Agarose Gel Electrophoresis	82
2.5.6.1	Agarose Gel Electrophoresis of DNA	82
2.5.6.2	Agarose Gel Electrophoresis of RNA	83
2.5.6.3	Molecular Weight Markers	83
2.5.7	Transfers to Membrane Filters	83
2.5.7.1	Southern Blotting (Southern, 1975)	83
2.5.7.2	Bidirectional Southern Blotting	84
2.5.7.3	Northern Blotting (Thomas, 1980)	85
2.5.8	Radioactive Labelling of DNA and Autoradiography	85
2.5.8.1	Radioactive Labelling of DNA by Random Priming	85
2.5.8.2	Measurement of Radioactivity Incorporated into DNA	86
2.5.8.3	Precipitation of Radiolabelled DNA Molecules	86
2.5.8.4	Autoradiography	86
2.5.8.5	RNA probes	87
2.5.9	DNA sequencing by the Dideoxynucleotide chain termination method	87
2.5.9.1	M13	88
2.5.9.2	pBluescript - single strand rescue	88
2.5.9.3	Double stranded templates	88
2.5.9.4	Sequencing reactions	89

2.5.9.5	Polyacrylamide gel electrophoresis for sequencing	91
2.6	Construction of genomic libraries from transposants K8 and A22	91
2.7	Polymerase chain reaction	92
2.8	Hybridisation of Northern Blots, Southern Blots and Plaque Lifts	93
2.8.1	Post-Hybridisation Washing	93
2.9	Radioactive Labelling of poly (A) ⁺ RNA by Reverse Transcription	94
2.9.1	Probing Southern Blots with Radiolabelled cDNA	94
2.10	In situ Hybridisation to Third Instar Polytene Chromosome Preparations with [3H]-dCTP labelled Probes	95
2.10.1	Preparation of Third Instar Polytene Chromosomes	95
2.10.2	Prehybridisation of Polytene Chromosome Preparations	96
2.10.3	Preparation of [³ H]-dCTP Labelled DNA Probes and Hybridisation	96
2.10.4	Washing Steps After Hybridisation	97
2.10.5	Autoradiography	97
2.10.6	Microscopy and Chromosome Staining	97
2.11	<i>in situ</i> hybridisation to polytene chromosome preparations with non-radioactively labelled probes	98
2.12	<i>In situ</i> Hybridisation To Whole Mounts of <i>Drosophila</i> Ovaries and Testes	99
2.12.1	Preparation of Tissues - Paraformaldehyde Fixation	99
2.12.2	Probe Labelling Using Digoxigenin - Non Radioactive Method	100
2.12.3	Pretreatment of Tissues	100
2.12.4	Hybridisation and Washing	101
2.12.5	Signal Detection	101

2.13	Detection of β -galactosidase expression patterns <i>in situ</i>	102
CHAPTER THREE		
RESULTS		
3.1	Screening Enhancer-trap lines for sex-specific β -galactosidase staining patterns	104
3.1.1	Preliminary molecular characterisation of "testis-specific" lines.	106
3.2	Screening K8 and A22 genomic libraries	107
3.3	Identifying the chromosomal locations of the insertions	108
3.4	Location of transcription units in the vicinity of the K8 and A22 insertions	108
3.4.1	Transposant K8	109
3.4.2	Transposant A22	110
CHAPTER FOUR		
4.1	Isolation of cDNA clones representing transcribed sequences within A 5.1	113
4.1.1	Mapping of cDNA clones ov.1 and ov.3 on to the cloned genomic region	113
4.2	A-5.1 includes part of a gene encoding multiple independently regulated transcripts	114
4.3	Isolation and characterisation of a cDNA clone representing the testis specific transcript	117
4.4	Isolation of genomic DNA encoding the testis-specific transcript	118

CHAPTER FIVE

5.1	Sequences of ov.1 and T1	120
5.1.1	T1 cDNA	120
5.1.2	ov.1 cDNA	120
5.2	Is ov.1 an authentic cDNA ?	121
5.3	The predicted products of T1 and Ov.1 show high similarity to human and mouse inosine 5' monophosphate dehydrogenases (IMPD)	124
5.4	Isolation of a cDNA encoding the full length <i>Drosophila</i> IMPD	125
5.5	Further investigation of the expression pattern of <i>Drosophila</i> IMPD	127
5.5.1	Northern analysis	127
5.5.2	Detection of IMPD transcripts <i>in situ</i>	128
5.5.2.1	Testis IMPD expression	129
5.5.2.2	Ovarian IMPD expression	130
5.5.2.3	Embryo in situs	131

CHAPTER SIX

6.1	Functions of IMPD	133
6.2	IMPD is implicated in tumorigenesis	133
6.3	<i>Drosophila</i> IMPD	134
6.4	Future developments	138
	References	143

CHAPTER ONE

INTRODUCTION

Introduction

On account of its unique amenability to genetic and molecular manipulation, *Drosophila melanogaster* has been highly informative in the study of higher metazoan development. Recently, seminal research by a number of groups, combining molecular technology with more traditional genetic approaches has realised the long standing goal of a molecular explanation of embryonic pattern formation (reviewed in Nüsslein-Volhard, 1991).

Sex-determination and sexual differentiation are other areas of developmental biology in which *Drosophila* research has greatly advanced our understanding in recent years. Sexual differentiation in a complex eukaryote involves a binary decision between diverse pathways of development and as such offers a paradigm for general processes of differentiation and has long been the focus of intensive research. In this section, I will describe the current model of *Drosophila* sex-determination, which has been developed from a synthesis of genetic and molecular data. Somatic and germline sex determination are considered separately since the mechanisms involved in these two cell types are different (reviewed in Slee and Bownes, 1991).

1.1 Somatic sex-determination

Somatic sex-determination is a cell autonomous process in which the primary signal is the ratio of X chromosomes to autosomes (X:A ratio) (Bridges, 1925 a and b). Male flies have one X chromosome and one Y chromosome whilst females have two X chromosomes. This primary signal is translated into the programme of sexual differentiation via a hierarchical series of regulatory genes, which were initially identified on account of their sex transforming mutant phenotypes. The principal genes of the hierarchy have since been cloned and their functions elucidated at the molecular level. The hierarchy is illustrated

schematically in Fig 1.1.

In females, the presence of two X chromosomes in conjunction with the maternally supplied products of the genes *daughterless* (*da*) (Cline, 1978), *Daughter-killer* (*Dk*) (Steinmann-Zwicky et.al., 1987) and *sans-fille* (*snf*) (Oliver and Mahowald, 1988; Steinmann-Zwicky, 1988) causes the activation of the key regulatory gene *Sex-lethal* (*Sxl*). In males, with one X chromosome *Sxl* remains inactive. Assessment of the X:A ratio appears to involve a balance between positively acting X-linked numerator elements and negatively acting autosomal numerator elements. Two known X-linked numerator elements are *sisterless-a* (*sis-a*) and *sisterless-b* (*sis-b*) (Cline, 1988). *sis-b* corresponds to the T4 locus of the *achaete-scute* complex (Torres and Sanchez, 1989) and encodes a putative transcription factor with a helix-loop-helix (H-L-H) dimerisation motif as does the *da* gene product (Murre et.al., 1989). One theory of X:A ratio assessment which accommodates the data envisages the formation of *sis-b/da* heterodimers which activate transcription from the *Sxl* locus in pre-cellular embryos (Parkhurst and Isch-Horowicz, 1990). The other element of the model is a conjectural autosomally encoded H-L-H protein which sequesters the *sis-b* or *da* products forming inactive heterodimers. Accordingly, in females with a double dose of *sis-b* the balance would favour the formation of the active dimer whilst in males, with one dose of *sis-b* insufficient *sis-b* product would be available for *Sxl* activation to occur. At best this could only provide a partial explanation of the mechanism involved since it does not accommodate the involvement of an additional X-linked numerator element, the gene *runt*, which does not encode an H-L-H protein (Duffy and Gergen, 1991).

As well as the appropriate X:A ratio, genetic studies have revealed that the activation of *Sxl* in female embryos requires a zygotically expressed product of the gene *snf* (Oliver and Mahowald, 1988; Steinmann-Zwicky, 1988).

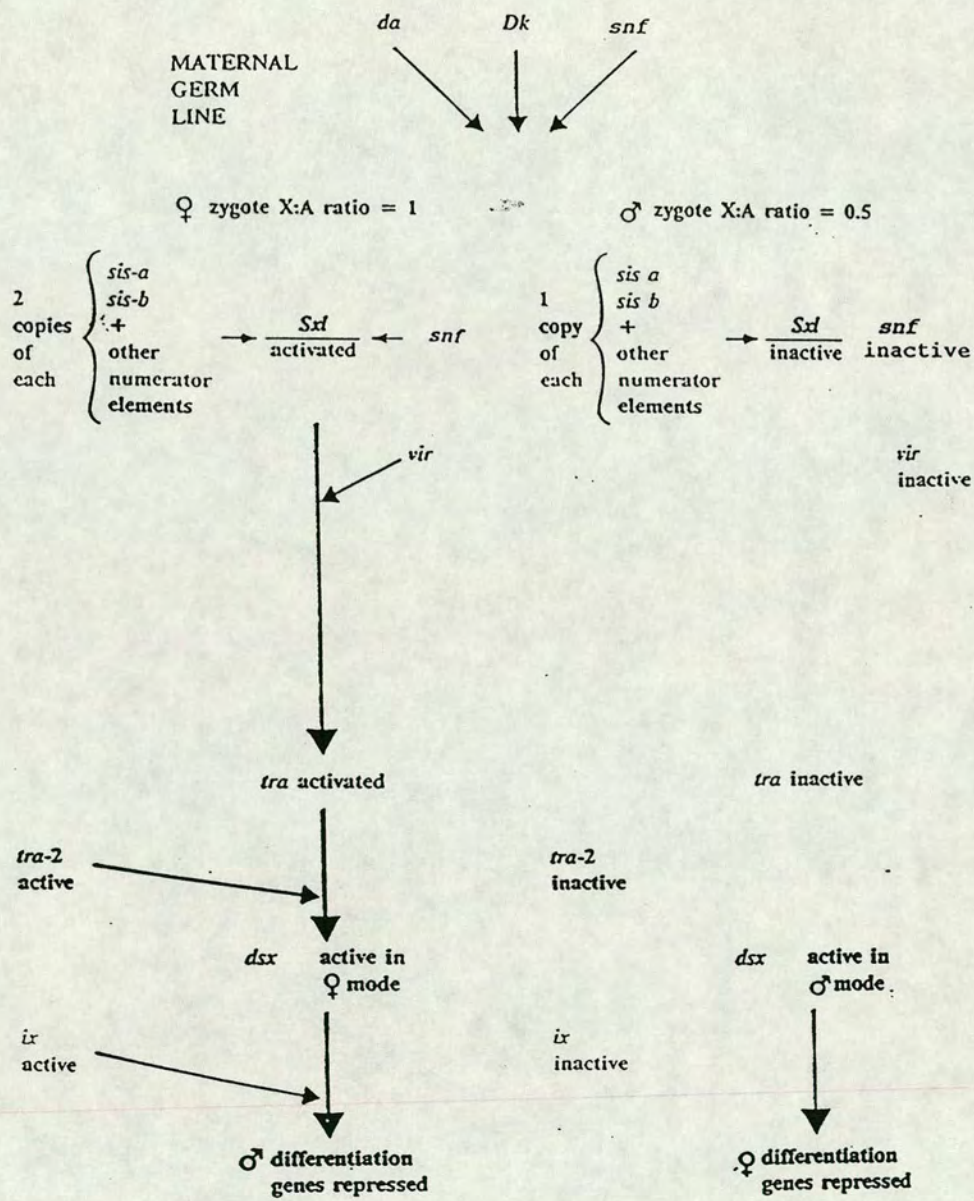
Following the initial female specific activation of

Figure 1.1

**The genetic hierarchy governing somatic
sex-determination in *Drosophila melanogaster***

In females, the presence of two X chromosomes and their associated numerator elements, in combination with a number of maternally and zygotically supplied gene products, leads to the activation of the key regulatory gene *Sxl*. The *Sxl* product acts through a branched hierarchy of additional genes to activate the female specific mode of expression of the bifunctional *dsx* gene and the female *dsx* product represses male differentiation genes, allowing female differentiation to occur.

In males, with one X chromosome and a single dose of numerator elements, *Sxl* remains inactive and the *dsx* gene is active in its default male specific mode, repressing female specific differentiation genes and permitting male differentiation to occur.



Sxl a post transcriptional regulatory mechanism comes into operation. Identical primary transcription products are produced in both sexes but an active *Sxl* protein is only produced in females, where the early *Sxl* product directs the excision of an exon from the later transcript, which contains an in frame stop codon and allows the production of a full length protein (Bell et.al., 1988). In males this exon is retained in the mature transcript truncating the open reading frame and resulting in an inactive protein. The active female specific protein establishes an autoregulatory loop by maintaining the female specific splicing pattern of the *Sxl* primary transcript. Once this loop is established the gene remains active in females independently of the X:A ratio (Bell et.al., 1991).

In addition to its autoregulatory function the female *Sxl* protein regulates the gene below it in the somatic sex-determination hierarchy, *transformer* (*tra*). This regulation is also at the level of differential RNA processing and requires an undetermined function encoded by the zygotically expressed gene, *virilizer* (*vir*) (Hilfiker and Nöthiger, 1991). In males, the *tra* transcript is spliced so as to include a stop codon which truncates the open reading frame and results in an inactive protein. In females, only a proportion of the *tra* primary transcript is spliced in this fashion and the remainder is spliced differently to exclude the stop codon and allow the expression of a functional gene product (Boggs et.al., 1987; Nagoshi et.al., 1988). Experiments have shown that this female specific splicing pattern results from the blocking of an alternative 3' splice acceptor by *Sxl* protein (Sosnowski et.al., 1989).

In females the *tra* product acts in conjunction with the product of another gene *transformer-2* (*tra-2*) to impose a female specific splicing and polyadenylation pattern on transcripts from the *doublesex* (*dsx*) gene (Nagoshi et.al., 1988; McKeown et.al., 1988). This regulation involves a positive control mechanism in that *tra/tra-2* proteins

promote the use of a female specific splice acceptor. Males represent the default state, where in the absence of *tra* protein an alternative 3' splice acceptor is utilised (Hedley and Maniatis, 1991; Hoshijima et.al., 1991).

The male and female *doublesex* transcripts both encode functional proteins, with common N-termini but different C-termini (Burtis and Baker, 1989) and the *dsx* gene acts negatively as a binary switch selecting either the male or female developmental pathway. The female product, probably in conjunction with the product of *intersex* (*ix*), acts to repress male differentiation, permitting female differentiation to proceed. Conversely, the male *dsx* product represses female differentiation and permits male differentiation (Baker and Ridge, 1980; McKeown et.al., 1988). The *dsx* proteins are thought to exercise their repressor function by binding *cis*-acting elements associated with the downstream target genes. Since *dsx* is the regulatory gene directly responsible for setting in train all the diverse processes of differentiation underlying the physiological and behavioural sexual dimorphism seen in adult flies it might be expected to act upon numerous target genes at various stages of development. Characterisation of as many of these target genes as possible would be an obvious next step towards a complete understanding of the sex-determination process.

1.1.1 The search for genes downstream of the regulatory hierarchy

One approach to the search for candidate target genes of the sex determination hierarchy is to use differential screening in order to identify genes with sex-specific expression patterns and this technique has been employed repeatedly, but with limited success. Although a substantial number of genes expressed exclusively in males or females have been isolated by this approach, (Barnett et.al., 1980; Di Benedetto et.al., 1987) only in the case of

the yolk protein genes has this expression been shown to be governed directly by the regulatory hierarchy (Bownes and Nöthiger, 1981). Possible reasons for the failure to isolate more genes in this class are discussed below.

1.1.2 The yolk protein genes

The three yolk protein genes (*yp1*, *yp2* and *yp3*) are expressed in the fat body and ovarian follicle cells of adult female flies (Bownes and Hames, 1978; Brennan *et.al.*, 1982). The fat body expression of these genes has been demonstrated to be continually dependent on the sex-determination hierarchy in experiments performed with chromosomally female flies homozygous for a temperature sensitive *tra-2* mutation (*tra-2^{ts}*) (Bownes and Nöthiger, 1981). At the restrictive temperature these flies develop as males whilst at the permissive temperature they develop as females (Belote and Baker, 1982 and 1983). Transcription of the *yp* genes in the female fat body can be switched on or off by shifting flies between the permissive and restrictive temperatures, indicating continuous dependence on a functional *tra-2* gene and by inference the other genes of the regulatory hierarchy. Indeed, the *cis*-acting elements mediating this *yp* gene regulation have been defined in P-element transformation experiments (Garabedian *et.al.*, 1986; Shirras and Bownes, 1987; Liddell and Bownes, 1991) and it has recently been shown that *dsx* proteins bind within the defined *cis*-acting regions mediating the fat body expression of *yp 1* and *yp 2* (Burtis *et.al.*, 1991).

1.1.3 Gonadal sex-specific gene expression

In contrast to the fat body expression, expression of the *yp* genes in the ovarian follicle cells is not sensitive to temperature shifts in adult *tra-2^{ts}* homozygotes (Bownes *et.al.*, 1990) and depends upon *cis*-acting sequences separate from the regulatory elements governing fat body

expression (Garabedian et.al., 1985; Logan et.al., 1989). These findings raised the possibility that sex-specific gene expression within sex-specific tissues may be irreversibly determined prior to adulthood and not continuously dependent on the regulatory hierarchy, a model which has been confirmed by subsequent research. All the genes isolated in differential screens on the basis of sex-specific expression, apart from the *yp* genes, are expressed exclusively in gonadal tissue and where tested the maintenance of this expression in adults has been shown to be independent of the hierarchy (Di Benedetto et.al., 1987; Chapman and Wolfner, 1988). However, temperature shifts at earlier developmental stages have revealed a temperature sensitive period during the late third larval instar when the later expression of these genes is determined (Chapman and Wolfner, 1988). Thus it appears that expression of *dsx* in the appropriate mode is necessary only to selectively trigger the development of the male or female genital primordium whilst gene expression within the gonad is controlled by tissue-specific factors (Wolfner, 1988).

There could be several reasons for the *yp* genes being the only *dsx* targets to have emerged from differential screens. Many of the genes *dsx* acts upon may be expressed only transiently during metamorphosis, as is presumably the case for its targets in the genital primordia, whilst other genes regulated by the hierarchy may not be expressed exclusively in males or females but may exhibit different spatial or temporal expression patterns in the two sexes. If such genes do not have an early vital function they might be detectable on the basis of a sex-specific mutant phenotype. Alternatively, the recently developed enhancer trap system, which is described in section 1.14, could provide a route to their isolation.

1.2 Germline sex-determination

As outlined in the preceding section, the same

regulatory hierarchy appears to govern, directly or indirectly, sex-specific gene expression in all somatic tissues of the fly. In the germline however, sex is determined by a different mechanism. This is best illustrated by experiments involving the transplantation of pole cells which are the progenitors of the germline and are the first cells to form in the *Drosophila* embryo. Pole cells from embryos mutant in any of the autosomal genes of the regulatory hierarchy differentiate according to their chromosomal sex if transplanted into a wild type embryo of matching sex, indicating that there is no cell autonomous requirement for *tra*, *tra-2* or *dsx* in germline sex-determination (Van Deusen, 1976; Marsh and Wieschaus, 1978; Schupbach, 1982). As well as demonstrating the dispensability of these autosomal genes, pole cell transplantations have revealed that germline sex is not determined in truly cell autonomous fashion. Mosaic studies indicate that somatic cells differentiate according to their chromosomal sex, regardless of the sex of the surrounding tissue but XX germ cells will enter spermatogenesis in a testis (Nöthiger et.al., 1989; Steinmann-Zwicky et.al., 1989). The influence of the somatic environment on germline sex determination is mediated by *Sxl*. Thus, the presence of the gain of function allele *Sxl*^{M#1} in XX pole cells allows them to enter oogenesis, even in a testis. Conversely, XX germ cells homozygous for a loss of function *Sxl* allele enter spermatogenesis, even in an ovary (Steinmann-Zwicky et.al., 1989). So it appears that *Sxl* activation in the germline may require an inductive influence from the female soma. Alternatively, the testicular soma may exercise a repressive influence on *Sxl*. Whichever is the case, the mechanism clearly contrasts with the situation in somatic tissues where *Sxl* activation is determined in a cell autonomous fashion by the X:A ratio. A number of other genes, of the so called "ovarian tumour" class, have been proposed to act in female germline sex-determination (Nöthiger et.al., 1989). These genes are

characterised by mutant phenotypes involving abnormal female germ cell proliferation, resulting in cysts which contain an excessive number of undifferentiated cells, resembling spermatocytes. Two of these genes are *snf* (which has been mentioned in relation to its role in somatic sex determination) and *ovarian tumour (otu)* (Steinhauer et.al., 1989). Tumours resulting from mutations at either of these two loci have been shown to express genes which are normally active only in male germ cells (Steinmann-Zwicky et.al., 1989 and G. Wei, B. Oliver and A.P. Mahowald unpublished in Pauli and Mahowald, 1990) and in the case of *snf* and another "ovarian tumour" gene *ovo* (Busson et.al., 1983) mutant alleles show genetic interactions with *Sxl* which lend support to their proposed roles in germline sex-determination (Steinmann-Zwicky, 1988; Oliver and Mahowald, 1988). Somatic recombination with loss of function *ovo* alleles has determined that the gene is required from the blastoderm stage onwards for survival of female germ cells (Oliver et.al., 1990). This finding is relevant to the earlier discussion of germline sex-determination since it implies that germ cell sex is determined around the same time as somatic sex i.e. at blastoderm, prior to the association of the germ cells with mesodermal cells. Therefore, the initial activation of *Sxl* in the female germline may not require an inductive influence from the ovarian environment, as has been suggested.

Recently another gene, *bag of marbles (bam)* has been identified, with a female mutant phenotype typical of the "ovarian tumour" class (McKearin and Spradling, 1990). However, this gene is distinguished by the fact that it is also required for normal male germ cell proliferation and therefore probably not directly involved in sex-determination. Instead, the authors propose that any sex-transforming effect of *bam* mutations on female gametogenesis is due to a breakdown of communication between soma and germline in the disorganised tumorous ovarioles.

Although there is no genetic evidence of a role in the male germline for any "ovarian tumour" class genes, like *bam* the gene *otu* is expressed in the testis as well as the ovary (Mulligan *et.al.*,1988) suggesting that it may have some subtle or redundant function in spermatogenesis.

1.3 Gonadogenesis/gametogenesis

Sexual dimorphism is most obvious in the reproductive organs of adult flies, where it extends from the gross morphology down to the levels of cell type composition and gene expression. This extreme dimorphism reflects the widely divergent functions of the male and female reproductive organs. In males, the testes and associated organs are responsible for the production of sperm and seminal fluid whilst in females the role of the reproductive organs is to produce oocytes, which are endowed with the maternally supplied gene products necessary for the early stages of embryogenesis. Any comprehensive picture of sexual differentiation must clearly include an explanation of the generation of the reproductive organs ,including the differentiation of the embryonic gonad and the events leading from the initial selection of male or female genital primordium to the mature accessory organs. Experimental evidence indicates that the sex-determination hierarchy acts during the third larval instar to select development of the appropriate genital primordium which forms the somatic tissues of the gonad. Subsequently, the differentiation of these tissues is independent of the hierarchy. Presumably, some cascade of regulatory loci must control the series of morphogenetic events leading to the mature structures but attempts to genetically dissect the process have been uninformative and may be hampered by pleiotropism in the genes involved. In the following sections, I will briefly describe the morphology and functions of the reproductive systems of males and females and the developmental processes leading

to their formation. Then I shall consider how genetic and molecular research has advanced our understanding in some areas, notably male and female gametogenesis.

1.4 Description of reproductive organs and gametogenesis in both sexes

1.4.1 Male

The male reproductive system (Fig 1.2A) consists of two coiled testes, each joined to a seminal vesicle by a short testicular duct or vas deferens. The paired seminal vesicles narrow and join to form the anterior ejaculatory duct which leads to the ejaculatory bulb from where sperm are conducted to the penis via the posterior ejaculatory duct. Two large accessory glands, or paragonia, empty into the anterior ejaculatory duct. They are responsible for the production of fluid necessary for the transmission of sperm. This fluid also contains sexual pheromones and factors which upon transfer to the female stimulate oviposition (Garcia-Bellido, 1964a.)

1.4.1.1 Spermatogenesis

(Reviewed in Lindsley and Tokuyasu, (1980) and Bownes and Dale, (1982)).

At the apex of each adult testis are situated 5-9 germ cells (Hardy et.al., 1979). A germ cell is estimated to divide approximately once every ten hours to produce another stem cell which remains attached to the apex and a primary spermatogonial cell which is surrounded by two somatically derived cyst cells. The descendants of the spermatogonial cell constitute a spermatogenic cyst and they remain enclosed by the cyst cells throughout their proliferation and differentiation into mature sperm. The single spermatogonial cell gives rise to 16 spermatocytes

Figure 1.2A

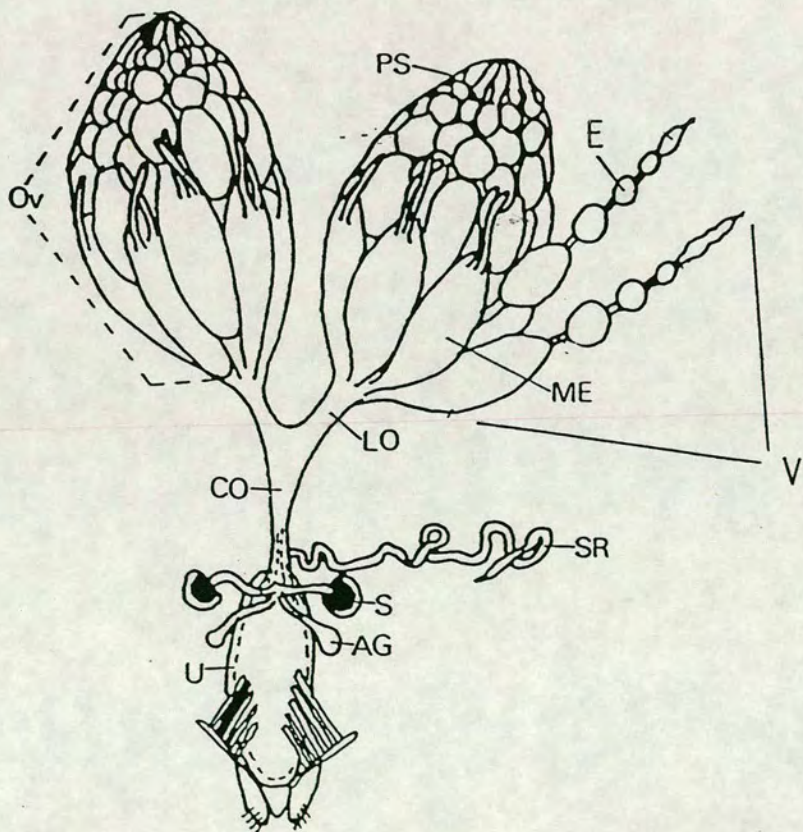
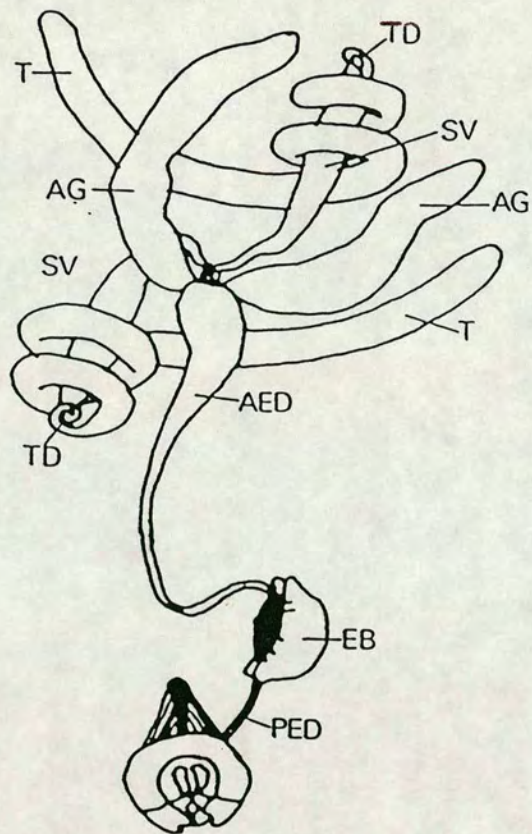
**Structure of the adult male reproductive system in
*Drosophila melanogaster***

T = testis
TD = testicular duct
AED = anterior ejaculatory duct
EB = ejaculatory bulb
PED = posterior ejaculatory duct
AG = accessory gland
SV = seminal vesicle

Figure 1.2B

**Structure of the adult female reproductive system in
*Drosophila melanogaster***

PS = peritoneal sheath
E = egg chamber
Ov = ovary
V = ovariole
ME = mature egg
LO = lateral oviduct
CO = common oviduct
SR = seminal receptacle
S = spermatheca
AG = accessory gland
U = uterus



through a series of 4 synchronous mitotic divisions. Cytokinesis is incomplete during these divisions so the progeny cells remain connected by intercellular bridges. Following the 4 mitoses the 16 cells undergo a 90 hour growth period during which they increase in volume 25 fold and mature into spermatids. Simultaneously, the cysts are displaced from a sub-apical position within the testis to the midpoint. As they descend the lumen of the testis, the cysts encounter bundles of elongated spermatids (see below) and are flattened against the testicular wall on the concave side of the testis. During this growth period the spermatocytes are active in transcription and translation. Since there is no further transcription after the spermatocyte stage (Gould-Somero and Holland, 1974), the RNA and protein produced at this stage must provide all the requirements for the subsequent meiotic divisions and the complex series of morphogenetic events which give rise to the mature sperm. The primary spermatocyte nuclei are distinguished from the nuclei of all other cell types by the presence of the Y chromosomal lampbrush loops which are cytologically visible extrusions of discrete portions of Y chromosome chromatin. These are thought to correspond to regions of high transcriptional activity (reviewed in Hess, 1980). At the midpoint of the testis, the spermatocytes undergo two meiotic divisions to generate 64 spermatids which again remain connected by intercellular bridges, due to incomplete cytokinesis.

1.4.1.2 Spermiogenesis

The differentiation of the spermatids into sperm involves a dramatic programme of morphological changes lasting 5-6 days, the salient features of which are described below.

a) Pre-elongation

After meiosis the mitochondria of each cell assemble into a bipartite laminate sphere called the nebenkern of about 7 μm diameter which later forms an important component of the sperm tail. The centriole fails to replicate in the second meiotic division and the single centriole present in each spermatid forms a structure called the basal body which later elongates to form the microtubular element of the sperm tail, the axoneme. Prior to the initiation of axoneme formation a section of the plasma membrane invaginates around the basal body to form a structure called the cilium and the acrosome is formed from the golgi apparatus.

b) Elongation

During the elongation period, which lasts about 35 hours the spermatids elongate to about 1.8mm and the cyst comes to occupy 80-90% of the length of the testis. Simultaneously, the cyst is displaced from the testis wall to the lumen. The axoneme grows to a length of about 1.7mm through the addition of tubulin molecules at its caudal end and the interleaved lamellate structures which constitute the nebenkern separate and extend into cylindrical rods, closely associated with the axoneme throughout its entire length. Paracrystalline material begins to be deposited in the larger of the two nebenkern derivatives. The nucleus is also dramatically transformed during the elongation phase, from a 6 μm sphere into a needle like structure of dimensions 10 μm X 0.1 μm . Following the completion of elongation there is a phase lasting about 33 hrs during which there is a further elaboration of the axoneme and the continued accumulation of paracrystalline material in one of the nebenkern derivatives.

c) Individualisation

The process of individualisation, lasting about 20 hrs involves the transformation of the bundle of syncytial spermatids into 64 individual sperm. During this period an "investment cone" forms around the head of each longitudinally aligned spermatid and moves caudally, carrying excess cytoplasm and organelles. In the wake of the advancing investment cones each spermatid is individually surrounded in plasma membrane and at the tail end of the sperm bundle the contents of each cone are deposited in the "waste bag".

d) Coiling

After the individualisation process one of the two somatically derived cyst cells, which is associated with the head region of the bundle, becomes attached to the terminal epithelium in the basal region of the testis and the sperm bundle is withdrawn from the apical portion of the testis and coiled. At the same time abnormal sperm are segregated from the rest and delivered to the waste bag. On completion of the coiling process the sperm are released from the cyst cells and travel to the seminal vesicle via the testicular duct.

1.4.2 Female

The adult female reproductive system (Fig 1.2B) comprises two ovaries, each attached to a lateral oviduct. The lateral oviducts unite to form the common oviduct which leads to the uterus, where mature eggs are fertilised with sperm stored in the seminal receptacle and spermathecae. Each ovary is covered by a network of muscle fibres called the peritoneal sheath and consists of 15-20 parallel ovarioles which are joined at their ends by terminal filaments.

1.4.2.1 Oogenesis

(Reviewed in Mahowald and Kambysellis (1980) and Bownes and Dale (1982)).

At the tip of each ovariole lie 1-4 germ cells called stem cells. Division of a stem cell generates a daughter stem cell and another cell termed the cytoblast, which undergoes four mitotic divisions, each followed by incomplete cytokinesis, to generate a cluster of 16 cystocytes joined by intercellular bridges. One cystocyte is destined to become the oocyte (or presumptive egg), while the remaining 15 cells become nurse cells, which transfer to the oocyte many of the requirements for early embryogenesis. The nurse cell / oocyte complex becomes surrounded by somatically derived pre-follicle cells, which later give rise to the various follicle cell types, one of which is responsible for secretion of the egg shell. Collectively, the oocyte, nurse cells and follicle cells are termed an egg chamber (Fig 1.1B). As an egg chamber matures, it enlarges and moves down the ovariole towards the oviduct, aided by muscular contraction, and since egg production is continuous, each ovariole contains a series of egg chambers at different stages of maturation. King *et.al* (1956) and King and Cummings (1969) have divided oogenesis into 14 stages upon which the following summary is based.

The stage 1 egg chamber consists of a 16 cell cyst surrounded by a monolayer of about 80 follicle cells and is located in the anterior-most region of the ovariole, called the germarium. During stages 2-5 the follicle cells proliferate to give a total of around 1200 and one of the anterior-most cystocytes undergoes meiosis and becomes morphologically distinguishable as the oocyte. This period is also characterised by the polyploidisation of the nurse cells in preparation for the heavy demands of the oocyte for nurse cell derived gene products, later in oogenesis. There is a gradient of polyploidisation such that the anterior-most cells, closest to the oocyte, undergo more

rounds of replication than the posterior cells and this process continues in later stages until the anterior cells have up to 2096 copies. The follicle cells also polyploidise but to a lesser extent. In stages 6 to 8 the follicle cells remain evenly distributed around the egg chamber and those surrounding the oocyte become columnar in shape but during stages 9-11 there is complex follicle cell migration and differentiation into several sub-types. Initially, most of the follicle cells move over the oocyte but subsequently clusters of 6-10 cells, called border cells, migrate between the nurse cells to the anterior surface of the oocyte. The columnar cells around the oocyte also differentiate and by stage 11 there are three distinct types; the anterior pole cells, the posterior pole cells and the columnar main body cells. About 80 squamous follicle cells are left surrounding the nurse cells. At about stage 8, the process of yolk deposition begins in which yolk proteins are endocytosed by the developing oocyte and this continues up to stage 10 and is accompanied by a dramatic increase in the volume of the oocyte. A proportion of the yolk protein is synthesised in the fat body and secreted into the haemolymph and the remainder is provided by the follicle cells surrounding the oocyte. The follicle cells are also responsible for the secretion of the vitelline membrane, during stages 9 and 10, which is the inner protective layer of the egg, and the completion of its formation coincides with the cessation of yolk protein uptake. The outer protective layer of the egg, or chorion, which has a complex laminate structure, is secreted by the columnar follicle cells and border cells between stages 11 and 14, after which the follicle cells degenerate. Concomitantly with the processes of yolk deposition and vitelline membrane formation, the nurse cells are active in the synthesis of material for provision to the developing oocyte. Prior to stage 10, RNA synthesis is intense and nurse cell volume increases, until at the end of stage 10B there is a rapid and massive movement of

nearly all the cytoplasm of the nurse chamber into the oocyte and the nurse cells degenerate. The nurse cell contribution includes large quantities of rRNA as well as the mRNA required for early embryonic protein synthesis. Other components are lipid droplets, which are delivered to the oocyte from about stage 8 onwards and various organelles. Over the entire period of oogenesis, the oocyte grows 100,000 fold and at stage 13 the nuclear envelope breaks down and the chromosomes move to the metaphase plate of meiosis 1 and arrest. Very little RNA is synthesised by the oocyte nucleus although a brief period of transcription has been detected autoradiographically during stages 9 and 10.

1.5 Development of the reproductive organs

The adult reproductive organs consist of the gonads and various accessory organs. Formation of the gonads begins early in embryogenesis while the accessory structures are largely derived from the genital imaginal disc, which only begins to develop during larval stages.

1.5.1 Gonadogenesis

The gonads are composed of germ-line and somatic cell types with distinct origins. The cells of the germ-line develop from the pole cells, which form at the extreme posterior of the embryo and the somatic cells are derived from mesodermal cells in the abdominal region. The two cell types come together through a complicated process of cell migration.

The pole cells, which are the progenitors of the germ-line are the first cells to form in the *Drosophila* embryo. The mature egg contains at its posterior, a specialised region of cytoplasm called the pole plasm, characterised by the presence of RNA rich polar granules (Mahowald, 1971). About 1.5 hours after fertilisation, following the 8th

mitotic division, the nuclei migrate to the surface of the embryo beneath the plasma membrane. Most of these nuclei are destined to form the syncytial blastoderm and yolk nuclei but about 10 are pinched off at the extreme posterior, where under the direction of localised pole cell determinants contained in the pole plasm (Illmensee and Mahowald, 1974), they acquire plasma membranes and form the pole cells, with a cytoplasm rich in polar granules. The pole cells undergo 2 or more mitotic divisions to produce a cluster of 40 - 60 cells. During gastrulation, these cells are carried into the embryo with the posterior midgut invagination. At about 7 hours post fertilisation, the pole cells move between the cells of the future midgut and come to lie between the midgut and the mesodermal layer (Sonnenblick, 1941). Here they become surrounded and intermingled with mesodermal cells to form the embryonic gonads. Only 50-60% of the pole cells are estimated to complete the migration to the presumptive gonads, the remainder either degenerating or being excreted through the anus (Underwood *et.al*, 1979). The developing gonads contain 5-13 germ cells and Sonnenblick (1941) has described a bimodal distribution with some embryos containing 5-7 and others 9-13. These are proposed to represent female and male embryos respectively, making this the earliest sign of sexual differentiation.

There is little known about the embryonic development of the gonad but by 16 hours a mesodermal sheath clearly demarcates it from other embryonic structures and it begins a steady process of growth (Sonnenblick, 1950). Between now and hatching, one or two mitoses of the germ cells are reported to occur so that by the first larval instar the testes contain 36-38 spermatogonia and ovaries contain 8-12 oogonia (Sonnenblick, 1941). By this stage, the testes are spherical, transparent bodies, much larger than the ovaries, providing a convenient means of distinguishing the sexes. Gonads of both sexes contain mesodermally derived somatic cells in addition to the germ cells. In the testis,

there are three somatic cell types; apical and interstitial cells at the anterior, which are later found at the apex of the mature testis and terminal cells at the posterior, which are progenitors of the terminal epithelium. The somatic component of the ovary appears undifferentiated in the larval stages but will later give rise to the follicle cells, the ovarian sheaths, the basal stalks and the calyx at the ovary posterior.

Data on the genetic regulation of gonadogenesis are sparse but a number of genes involved in germ cell differentiation have been identified. These include the seven so called "grandchildless" genes, whose maternally provided products have an unspecified role in the formation of the pole-plasm and pole cells in the early embryo. Embryos derived from mothers homozygous for mutations at any of these loci fail to form polar granules or pole cells. In addition, these embryos exhibit posterior pattern defects and the genes will be further discussed in a later section in the context of their function in posterior pattern formation.

1.5.2 Formation of the adult reproductive system

Apart from the aforementioned, all the somatic structures of the male and female reproductive systems are derived from the genital imaginal disc, which comprises separate primordia for the male and female genitalia and a single anal primordium (Nothiger *et.al*, 1977). The presumptive disc cells first become visible on histological sections in 22-25 hour embryos (Laugé, 1967 in Laugé, 1982). During the embryonic and larval periods the genital discs grow through a combination of cell division and cell enlargement. There is no histological evidence of sexual dimorphism until the end of the second larval instar and this becomes more evident during the third instar, when the male disc stands more upright and the female disc lies flat

on the larval epithelium. After pupariation there is a shift in the pattern of disc development from cell multiplication to cell differentiation and the complex morphogenetic process termed eversion, leading to formation of the genitalia and analia. In the male, early in pupal development, at about 31 hours after puparium formation (APF), two anterior projections of the genital disc, the presumptive seminal vesicles, join the posterior extremities of the testes which are by now ellipsoidal in shape. The epithelial linings become continuous at the point of junction and pigment cells from the testes migrate over the seminal vesicles. This fusion of testes and disc is thought to stimulate testis elongation since failure to join is correlated with failure of elongation. However, it is not a prerequisite for the production of motile sperm (Stern and Hadorn, 1938) and larval testes transplanted into adult females are capable of normal spermatogenesis (Garcia Bellido, 1964b in Lindsley and Tokuyasu, 1980). Over the course of pupation, the genital disc develops the remaining components of the reproductive system including the accessory glands, ejaculatory duct and external genitalia. A broadly analogous series of events occurs in the female, where the first contact between the gonad and genital disc derived structures occurs at about 36 hrs APF when the lateral oviducts join the ovaries and over the following six hours this union is further established, through the migration of oviduct cells over the ovary surface and the eventual formation of the calyx at the point of junction, with cell contributions from both structures. The development of the lateral oviduct appears to be independent of the presence of the ovary but tissue specific recognition is involved in their connection (Pantelouris, 1955). Interestingly, the association between gonad and genital disc does not seem to be sex-specific, since larval ovaries transplanted into male hosts are often found connected to the male reproductive tract (Kambyzellis, 1968).

1.6 Mosaic analysis in the study of development

Cell lineage relationships between different elements of the reproductive system can be determined by clonal analysis, which can reveal when compartmentalisation of cells according to developmental fate occurs and also the spatial relationship between different compartments. The technique relies on the generation of genetic mosaics, in which the contribution of phenotypically distinguishable clones of cells generated early in development to different adult structures, can be assessed. The underlying assumption is that the frequency with which two distinct pattern elements are encompassed by a single clone is a function of the proximity of their primordia and the developmental stage at which the clonal restriction occurs. The mosaics studied are usually gynandromorphs which are XX \ XO mosaics consisting of patches of male and female tissue resulting from sporadic loss of an unstable ring X chromosome during syncytial cleavage. The use of such mosaics in developmental studies is reviewed in Hall *et.al.*, 1976. Alternatively, mosaics may be artificially generated by somatic recombination. In the latter case, clones homozygous for a recessive marker mutation such as *yellow* are produced in heterozygotes by X-ray irradiation.

Examination of gynandromorphs, in which elements of the male **and** female genitalia can develop, established that there are completely separate male and female primordia rather than a single primordium as previously supposed. The existence of distinct primordia was conclusively evidenced by the presence in some gynandromorphs of two complete sets of genitalia: one male, one female. (Nothiger, 1977).

Analysis of mosaics resulting from X chromosome loss or somatic recombination induced at blastoderm, and progressively later developmental stages, has been informative with respect to the spatial relationship between the anal and genital primordia. It has also afforded insight into the differentiation of cells in the

developing primordia (Dubendorfer, 1977; Schupbach *et.al*, 1978). The results indicated that as early as blastoderm, the precursors of the genitalia and analia formed clonally separate cell populations. However there is no clonal restriction between cells contributing to the right or left sides of the genitalia and analia until at least the second larval instar.

The relative distances between primordia or between precursors of different pattern elements within a primordium, can be estimated from mosaic analysis according to the frequency with which their adult derivatives are included within a single clone. The unit of measurement is the sturt. Two primordia are said to be one sturt apart when their adult derivatives are of different genotypes in 1% of the mosaics examined. Schupbach (1978) has applied this method to the production of detailed fate maps of the male and female external genitalia and analia. Similarly, mosaic studies employing gynandromorphs have revealed that the nuclei destined to form pole cells lie approximately 25 sturts posterior and slightly dorsal to the precursors of the genital disc (Gehring *et.al*, 1976).

1.7 Surgical techniques

Fate maps can also be deduced from surgical experimentation. Imaginal discs are removed from late third instar larvae and surgically treated by UV microbeam radiation or mechanical fragmentation. The developmental potential of the surviving tissue can then be assessed by transplantation into host larvae. Two dimensional fate maps of the male and female genital discs have been derived by these means (Bryant and Hsei, 1977; Littlefield and Bryant, 1979).

1.8 Genetic regulation of development of the reproductive system

Mutational analysis has yielded little information concerning the genetic regulation of the differentiation of the genital disc, presumably because most of the genes involved will have other vital functions. Initial selection of the male or female genital primordium is known to be under the control of the somatic sex-determination gene hierarchy as discussed previously and initiation of the appropriate morphogenetic process presumably involves the local activity of homeotic genes. In order to identify genes further downstream in the pathway novel techniques such as the enhancer-trap (see below) may prove useful.

1.9 Genetic and molecular analysis of spermatogenesis

Drosophila spermatogenesis represents an attractive system in which to study the genetic control of morphogenesis since it involves only a single cell lineage and consists of a series of well defined morphogenetic changes. Also, many of the genes with a role in spermatogenesis can be conveniently identified in simple screens for a male sterile phenotype. The complexity of the system is reflected in its susceptibility to disruption by mutations in a large number of genes. It has been estimated from experimental data, that the number of genes mutable to male sterility is equivalent to 15 % of the number of genes mutable to lethality. Furthermore, studies with temperature-sensitive lethal mutations indicate that up to one third of the lethally mutable loci are also crucial to male fertility leading to an estimate of the number genes involved as high as 1750 (Discussed in Lindsley and Tokayasu, 1980). However, there is also evidence that there may be considerable overlap in the groups of genes identified by male sterile or lethal mutations and that many male sterile mutations are in fact weak alleles of

vital, pleiotropic loci (discussed in Gönczy et.al., 1992). The selective effect on spermatogenesis of such alleles may reflect a particular sensitivity of the spermatogenic process to perturbations of general cellular functions.

Cytological examination of male sterile mutants can reveal the nature of the lesion concerned, allowing individual genes to be correlated with specific steps in spermatogenesis. However several authors have emphasised the need for caution in inferring the wild type functions of genes from the observable defects in spermatogenesis since the primary lesions in many mutants can give rise to a cascade of subsequent abnormalities (Lifschytz and Hareven, 1977; Hackstein, 1990). It is therefore important to determine the earliest deviation from normal development in every case. In one screen of the X chromosome the authors examined 600 male sterile mutants for cytological abnormalities in primary and secondary spermatocytes (Lifschytz and Hareven, 1977). Several genes were identified whose wild type products were required for the normal execution of meiosis. For example, ms(1)516 was shown to affect the bipolarity of the second meiotic spindle. Mutants possessed normal first division spermatocyte cysts but in the second meiotic division both centrioles located to the same pole resulting in monoastral spindles. In a number of other mutants exhibiting a failure of meiosis, defects were first noticeable at earlier, pre-meiotic developmental stages. For example, ms(1)401 caused various abnormalities of the primary spermatocyte including abnormal invaginations of the nuclear membrane and small fragmented nuclei and three mutations were isolated which caused premature mitochondrial aggregation, at the primary spermatocyte stage, an event which normally occurs during spermiogenesis. Interestingly, mutations in the latter class were pleiotropic and caused behavioural defects, possibly a consequence of a general defect in mitochondrial or microtubular function. All these meiotic mutants permitted development to proceed as far as spermatid

elongation. However, by this stage they all exhibited similar non-specific morphological disorganisation which gave no clues to the diversity of these genes' early functions.

By following the course of spermatogenesis in the various mutants, it was possible to draw certain conclusions regarding the interdependence of the genetic control of different aspects of the developmental programme. Nebenkerne formation is evidently independent of meiosis since it can precede meiosis in some mutants and nebenkerne elongation is not structurally dependent on the axoneme because in the same mutants, there is elongation without prior attachment to the axoneme. Hackstein *et.al.* (1990) has also reported mutations which uncouple separate pathways within the global programme, in a screen of *Drosophila hydei*. He was able to demonstrate that neither the normal number of gonial divisions nor the normal unfolding of the Y chromosome lampbrush loops are necessary for the shaping of the primary spermatocyte nucleus. Other mutants showed that the flagellum and spermatid nucleus can differentiate in the absence of a nebenkerne.

1.9.1 The Y chromosome

The sole function of the Y chromosome appears to be in spermatogenesis, since the only detectable consequence of a lack of this chromosome in males is sterility (Bridges, 1916 in Lindsley and Tokuyasu, 1980) and the fact that it is dispensible for viability has simplified its genetic analysis. By means of complementation tests with a series of X-ray induced male sterile Y chromosomes, Brousseau (1960) identified seven distinct chromosomal regions which were necessary for fertility. Subsequently, Kennison (1979 in Bownes and Dale, 1982) was able to confirm the existence of six of these so called fertility factors and assign them to sites on the long and short arms, by constructing a set of deficiencies for the Y chromosome. Although it normally

appears entirely heterochromatic, during the primary spermatocyte stage the Y chromosome forms lampbrush loops, which are generally considered to be a manifestation of transcriptional activity (Hess and Meyer, 1968). Because its lampbrush loops are very pronounced, *Drosophila hydei* is often preferred for the study of these structures and it has been possible to demonstrate in this species, that each of the genetically identified fertility factors corresponds to a particular lampbrush loop (Hess, 1980). In *D. melanogaster* it has only been possible to demonstrate such a relationship for three of the six fertility factors (Bonaccorsi et.al., 1988).

In addition to the six fertility factors, whose functions in spermatogenesis are poorly defined, two other genetic loci have been identified on the Y chromosome (discussed in Ashburner, 1989). One of these is a *bobbed* (*bb*) allele on the short arm which encodes ribosomal RNA and is involved in the organisation of the nucleolus. The other is the *stellate* locus, so called because it is deficiency for this locus which causes the appearance of the large star shaped crystals in the primary spermatocyte nuclei of XO males. The function of this gene is obscure, but it is known that in its absence homologous sequences on the X chromosome overproduce a 17kd protein which aggregates to form the observed crystals.

1.9.2 Molecular studies in spermatogenesis

In recent years, the application of molecular genetic approaches has resulted in the isolation of a number of genes with functions in spermatogenesis and provided insight into the nature of the gene products, patterns of expression and mechanisms of gene regulation. A notable example is the gene *mst* 87F (formerly *mst*(3)gl-9) which was isolated in a differential screen and found to be expressed exclusively in the male germline (Schäfer, 1986). The predicted protein product of this gene contains only 55

amino acids and consists predominantly of repeats of cysteine, glycine and proline. Its function is uncertain but mammalian spermtails are known to contain dense fibres associated with the axoneme, which are rich in cysteine and proline and the authors speculate that the *mst* 87F product may be a component of analogous structures in *Drosophila* sperm (Kuhn *et.al.*, 1988). In Northern analysis, transcripts from the gene were first detected in third instar larvae and become more abundant in the pupal stage. P-element transformation experiments with constructs fusing portions of 5' sequence to a reporter gene, determined that all the *cis*-acting sequences necessary for correct transcriptional regulation were contained within a fragment including 100 bp of 5' flanking DNA and 102 bp of untranslated leader sequence (Kuhn *et.al.*, 1988).

As stated previously, all transcription ceases well before meiosis in *Drosophila* spermatogenesis. It had been thought that this could reflect the inappropriateness of the imposition of selective pressure on the genetically heterogeneous haploid gametes. However, in view of the syncytial nature of the sperm bundle and the fact that all sperm effectively share a common gene pool up to the stage of individualisation, this principle should not apply, and indeed post-meiotic transcription is known to occur in mammalian spermatogenesis at the syncytial stage (Geremia *et.al.*, 1978). In any case, the situation in *Drosophila* means that many transcripts must be stored for long periods prior to translation and that post-transcriptional regulatory mechanisms must be important in the later stages of spermatogenesis. Immunofluorescence indicated there was no translation of the *mst* 87F mRNA until very late stages of spermatogenesis, about 3 days after the end of transcription and the *cis*-element mediating this translational repression was located to a 24bp sequence in the untranslated portion of the mRNA and named the translational control element (TCE). Sequence comparison with six other co-regulated genes which form a gene family

with *mst* 87F identified a highly conserved block of 12 nucleotides within the 24bp sequence, which was present at an identical position in all seven genes and may represent a general translational control element. Other genes expressed in testes but not under translational control, including β 2 tubulin (see below) were found to lack this element (Schäfer et.al, 1990)

Upon translational recruitment, the *mst* 87F poly(A) tail is extended from 140 to 380 nt and this modification is also directed by the TCE. This close coupling of translational repression and secondary polyadenylation led the authors to propose that a common trans acting protein may mediate both processes. Alteration of poly(A) tail length has been observed in several other developmental systems involving translational control although its functional significance is unclear (for review see Jackson and Standart, 1990). Evidence for another possible regulatory role of polyadenylation in *Drosophila* spermatogenesis has come from the molecular analysis of *janus A*, another gene transcribed in primary spermatocytes. This gene produces transcripts of identical coding capacity in the germlines of males and females, but the poly (A) tail of the male transcript is on average 180 bases longer, which may increase RNA stability during an extended storage period (Yanicostas et.al., 1989). In view of the translational control operating in spermatogenesis, it is interesting that the posterior group gene *vasa*, whose product shows homology to the human elongation factor eIF-4A and is believed to regulate translation in the early embryo, is expressed in the adult male germline (Lasko and Ashburner, 1988). However, as already stated *vasa* does not appear to have any essential role in spermatogenesis. One gene with an established role in RNA processing which is expressed during spermatogenesis, is *tra-2*. The *tra-2* product has a key role in somatic sex determination since it is responsible for directing female specific splicing of transcripts from the *dsx* gene (see section 1.1). The

mechanism appears to involve the binding of *tra-2* protein to a defined sequence within the female-specific exon of the *dsx* pre-mRNA, which then promotes the use of the female specific 3' splice site and polyadenylation site (Hedley and Maniatis, 1991). Molecular analysis has revealed that *tra-2* gives rise to multiple transcripts through differential splicing and promoter usage, all of which encode proteins with a putative RNA binding domain. Two of these transcripts are found in the somatic tissues of both sexes and in the female germline (Amrein et.al., 1990 ; Mattox et.al., 1990) The role, if any, of the male somatic and female germline transcripts is unclear since genetic analysis has identified no *tra-2* function in these tissues. Since *tra-2* seems to act in conjunction with a second gene, *tra*, in the regulation of female *dsx* expression and *tra* is not expressed in males, male somatic *tra-2* expression may indeed be non functional. *tra-2* is known to be required in the male germline for fertility and loss of function alleles have been shown to disrupt the later stages of spermatogenesis (Watanabe, 1975; Belote and Baker, 1983). Two alternatively spliced *tra-2* transcripts have been detected in the male germline, which differ in the presence or absence of a single intron, termed M1. Remarkably, this differential splicing is regulated by the *tra-2* product itself, in a mechanism reminiscent of *Sxl* autoregulation (see earlier). Apparently the *tra-2* product acts in the male germline to repress splicing of the M1 intron in 60-70% of transcripts (Amrein et.al., 1990; Mattox et.al., 1990; Mattox and Baker, 1991). Although the function of the *tra-2* products in the male germline remains to be established, it seems likely to involve some interaction with RNA, perhaps in the regulation of mRNA processing or stability. The observed effects of a *tra-2* mutation on the expression of the gene *exu*, support the idea that *tra-2* is involved in RNA processing in this tissue. The *exu* gene is known to function in spermatogenesis as well as embryonic pattern formation (Hazelrigg, 1990) and it has recently

been shown that a *tra-2* mutation (*tra-2^B*) disrupts the splicing and or polyadenylation of *exu* transcripts in the male germline (T.Hazelrigg, pers. comm. in Mattox and Baker, 1991). However it remains to be determined if this reflects a direct role of *tra-2* in *exu* RNA processing and whether or not *tra-2* has other functions in spermatogenesis.

1.9.3 The $\beta 2$ tubulin gene

The $\beta 2$ tubulin gene is expressed exclusively during spermatogenesis and encodes the major β tubulin isotype in the cytoskeleton of mature primary spermatocytes, in the meiotic spindles and in the sperm axoneme (Kemphues and Kaufmann, 1982; Kaltschmidt *et.al.*, 1991). This requires a high level of transcription and P-element transformation experiments with a $\beta 2$ tubulin \ lacZ fusion gene flanked by various amounts of upstream DNA, revealed that a 14bp sequence, B2 upstream element 1 ($\beta 2$ UE1) located at -51 from the transcription initiation site was sufficient to direct tissue and temporal specificity of transcription whilst a second 5bp element at -32 had a positive quantitative effect (Michiels *et.al.*, 1989). $\beta 2$ UE1 can direct male germline specific transcription from a truncated hsp70 promoter but it does not behave like a classical enhancer element since it shows a high degree of position dependence. Both $\beta 2$ UE1 and the downstream element exist at identical positions upstream of the *D.hydei* $\beta 2$ tubulin promoter, although flanking sequences have diverged significantly, indicating conservation of the transcriptional regulatory mechanism between these species. However, no such element is found upstream of *mst 87F* or other genes with testis expression that have been examined, implying that these genes are under separate transcriptional regulatory mechanism(s) (Michiels *et.al.*, 1989).

In future, as the molecular analysis of genes

expressed in spermatogenesis proceeds and *cis*-acting elements mediating transcriptional and post transcriptional regulation are defined, it should be possible to isolate the genes encoding the trans-acting regulatory proteins, perhaps by screening lambda expression libraries with the target sequences, as was achieved with a mammalian enhancer binding protein (Singh *et.al.*, 1988). Furthermore, once such regulatory proteins are available they may provide a means to the isolation of additional downstream target genes by a recently demonstrated PCR \ immunoprecipitation based method (Kinzler and Vogelstein, 1989). In the case of the *mst* 87F and β 2 tubulin genes, preliminary work towards the isolation of regulatory proteins is underway and there are many other cloned genes expressed in spermatogenesis whose regulation remains to be characterised. These include heat shock genes (Krawczyk *et.al.*, 1987; Allen *et.al.*, 1988), histone genes (Büsslinger *et.al.*, 1985; Cole *et.al.*, 1986) and several genes with unidentified products (e.g. Schäfer, 1986 ; Yanicostas *et.al.*, 1989 ; Schulz and Butler, 1989).

In conjunction with the molecular analyses, the generation of mutations *in vivo* may help to elucidate the function of cloned genes, employing either conventional mutagenesis or the more rapid PCR based system of site directed mutagenesis (Kaiser and Goodwin, 1990).

1.10 Genetic and molecular analysis of oogenesis

Oogenesis has been intensively studied at the genetic level and by screening for a female sterile phenotype, numerous genes involved in the regulation and execution of the process have been identified (e.g. Bakken, 1973; Gans *et.al.*, 1975; Mohler, 1977; Schupbach and Wieschaus, 1989). As in the case of spermatogenesis, the proportion of the genome involved in oogenesis is high and many of the genes involved may not be accessible by mutagenesis. Through the generation of mosaics in which the somatic cells of the fly

are wild type while the germline cells (oocyte and nurse cells) are homozygous for lethal mutations, it has been estimated that at least 75% of the genes required for viability are also required in the germline for oogenesis (Garcia-Bellido and Robins, 1983; Perrimon *et.al.*, 1984).

Female sterile mutations are conveniently classified according to the aspect of oogenesis which is affected. Class 1 comprises mutations which result in abortive egg production or the production of abnormal eggs which cannot be fertilised, whilst class 2 consists of the so-called "maternal effect" mutations, which perturb early embryonic development.

Oogenesis is a complex process requiring the concerted activities of the germ line cells and the various somatic tissues of the ovary and accessory organs. In addition, there are functions crucial to oogenesis which are provided by somatic tissues outwith the reproductive organs. Techniques are available which allow further subdivision of class 1 mutations according to which tissue or cell type is directly affected.

1.10.1 Ovary transplantation

Ovaries of mutant females can be transplanted into wild type larvae, where a proportion will connect with the duct system during metamorphosis. If suitable genetic markers are employed, examination of the progeny of the host will reveal if the mutant ovary is functioning normally in the wild type background. Failure to function infers that there is an ovary autonomous defect whilst a restoration of normal function indicates a defect outwith the ovary.

1.10.2 Pole cell transplantation and somatic recombination

The ovary includes both somatic and germ line cell types and there are two techniques which can be employed to

establish whether an ovary autonomous mutation affects the germ line or somatic cells. The first relies on the same rationale as the ovary transplantation experiments and involves transplantation of genetically marked pole cells between wild type and mutant embryos (Wieschaus and Szabad, 1979). Germ cell autonomous mutations are not rescued in a wild type host but mutations affecting somatic cells of the ovary are rescued and progeny are seen to be derived from the mutant pole cells. An alternative approach to determining which ovarian cell type harbours the defect is to generate germ line clones simultaneously homozygous for a visible marker and a linked female sterile mutation, by X-ray induced mitotic recombination in larvae heterozygous for the two mutations. The germline autonomous mutation *fs(1)k10* (Wieschaus et.al., 1981) has proved a useful marker for this purpose. Stem cells homozygous for *fs(1)k10* give rise to eggs with an obvious abnormal accumulation of chorionic material at the anterior pole and if the linked female sterile mutation is also germline autonomous then a proportion of *k10* eggs will also display the second phenotype. Otherwise, the mutation under test may be inferred to affect somatic cell function.

1.11 Ovary non-autonomous mutations

The primary lesion in ovary non-autonomous mutations which disrupt oogenesis may be a specific effect on another organ of the reproductive system or a constitutive defect which leads to a general impairment of cellular function. Some mutations in the latter category will be hypomorphic alleles of genes with vital functions. The best characterised ovary non-autonomous mutations affect the accumulation of yolk proteins in the developing oocyte (vitellogenesis) by disrupting the hormonal regulation of this process. One of the hormones involved in this regulation is juvenile hormone (Gavin and Williamson, 1976; Bownes et.al., 1983) and two examples of mutations in this

category are *apterous* (Gavin and Williamson, 1976) and *fs(2)A18* (Mahowald and Kambyzellis, 1980) both of which can be rescued by application of a juvenile hormone analogue to the mutant abdomen (Gavin and Williamson, 1976; Kambyzellis and Gelti-Douka, 1974). Moreover, *apterous* mutants were found to have a defective corpus allatum, which is the site of juvenile hormone secretion in the adult (Postlethwaite and Handler, 1978).

Detailed microscopic analysis of female sterile mutants can be highly informative with respect to the nature of the lesion, allowing inferences to be drawn regarding the wild type function of the gene concerned. In this way, mutations have been identified which affect a range of cell types and developmental events during oogenesis. For example the mutation *female sterile(2)B* (Yarger and King, 1971) disrupts the cystocyte divisions such that mitosis proceeds to completion and no intercellular bridges are formed. Two mutations which interfere with nurse cell function are *chickadee* and *kelch* (Cooley et.al., 1990), both of which result in undersized oocytes which remain attached to persistent nurse cells. The pleiotropic mutation *ocelliless* (Spradling et.al., 1979) affects chorion synthesis. Homozygous *ocelliless* females produce eggs which are greatly deficient in the levels of several major chorion proteins.

For the reasons which were discussed in relation to the genetic analysis of spermatogenesis, the information obtainable from such morphological analysis of female sterile mutants is limited. In the endeavour to understand the mechanisms controlling key events in the process of oogenesis, such as the differentiation of the presumptive oocyte from the 15 nurse cells or the interactions between the various germline and somatically derived cell types, a molecular approach will undoubtedly prove profitable. Presumably, intercellular communication and signal transduction pathways will be important elements, as has been demonstrated through a combination of genetic and

molecular research in the processes of *Drosophila* eye development and neurogenesis (reviewed in Hafer and Basler, 1991 and Artavanis-Tsakanos and Simpson, 1991).

1.12 Molecular analyses

Some progress has already been made towards a molecular understanding of oogenesis. A number of structural genes encoding major components of the mature egg have been cloned, facilitating the analysis of their expression and regulation at the molecular level. These include the yolk protein and chorion protein genes.

1.12.1 The yolk protein genes

The yolk proteins are expressed in the fat body and ovarian follicle cells of adult females (Bownes and Hames, 1978 ; Brennan *et.al.*, 1982) and endocytosed by the developing oocyte. The yolk protein genes were isolated in a differential screen by Barnett *et.al* (1980). Three distinct groups of clones hybridising to two sites on the X chromosome were identified and hybrid-arrested translation was used to demonstrate that each group encoded one of the three yolk polypeptides. Yolk protein genes 1 and 2 (*yp1* and *yp2*) are located at position 8F-9A on the polytene chromosome map and yolk protein gene 3 (*yp3*) is located more than 1000kb away at 12-BC. *yp1* and 2 are divergently transcribed and separated by an intergenic region of 1.2 kb (Hung *et.al*, 1982). P-element transformations with constructs comprising various portions of the intergenic region fused to a reporter gene, have identified two separate enhancers mediating the fat body and ovarian transcription of these genes. Transcription in the fat body is controlled by a 125bp element lying between -196 and -321 of the *yp1* cap site which acts in the position and orientation independent manner typical of an enhancer element (Garabedian *et.al*, 1986). The fat body

expression of the *yp* genes has been shown to be continuously dependent on the sex-determination hierarchy (see earlier) and it has recently been shown that *dsx* proteins bind to consensus sequences within the fat body enhancer, implicating the *dsx* product directly in the transcriptional regulation (Burtis *et.al.*, 1991).

Ovarian follicle cell expression of *yp1* and 2 depends on a combination of multiple elements separate from the fat body enhancer. P-element transformation experiments have identified DNA between -159 and -340 of the *yp2* cap site as necessary for ovarian expression of *yp1* and 2 (Garabedian *et.al.*, 1985; Logan *et.al.*, 1989). It is intriguing however, that this element was shown to act in an orientation dependent manner on a heterologous promoter since it is required to act bidirectionally on *yp1* and 2 *in vivo*. A second region has been identified within the first exon of *yp2* which increases *yp1* transcript levels in ovaries but does not confer ovarian expression from a heterologous promoter like a true enhancer element. An indication of further complexity in ovarian regulation has come from the isolation of a DNA binding protein with a very high affinity for a 31bp sequence 148bp downstream of the *yp1* transcription initiation site which is necessary for normal *in vivo* levels of *yp1* RNA (Logan *et.al.*, 1989).

As well as being under stage, tissue and sex-specific regulation, *yp* gene expression is modulated by the steroid hormone 20-hydroxyecdysone (Postlethwaite *et.al.*, 1980; Bownes *et.al.*, 1983) and nutritional status (Bownes and Blair, 1986). *Cis*-acting elements mediating the nutritional control have been localised to within 890bp 5' of *yp1*, whilst the ecdysone response depends on DNA outwith this region (Shirras and Bownes, 1987), and probably 3' of *yp1* or *yp2* (Shirras, Bownes and Wensink, personal communication). Analysis of *yp3* regulation is at a preliminary stage but it has been shown that the regulatory sequences governing fat body and ovarian expression lie within 706 bp 5' of the coding sequence (Liddell and

Bownes, 1991).

1.12.2 The chorion protein genes

The family of genes encoding the chorion proteins was also isolated in a differential screen (Spradling and Mahowald, 1980). Hybrid arrested translation demonstrated that the cDNA clones selected in the screen encoded three proteins corresponding in size to the three major chorion polypeptides and belonged to two groups hybridising to chromosomal locations 66D 15 and 7E11-7F12, the predetermined positions of gene clusters encoding the major eggshell proteins. Northern hybridisations with chorion protein gene probes demonstrated that transcription of the X linked genes begins at stage 11 of oogenesis whereas the 3rd chromosome cluster is transcribed at stages 13 and 14. However, within each cluster individual genes have unique mRNA accumulation profiles (Griffin-Shea *et.al.*, 1982; Parks and Spradling, 1987). P-element transformations with constructs fusing various amounts of upstream DNA to an Alcohol Dehydrogenase (ADH) reporter gene have demonstrated that for all the genes tested the *cis* acting elements necessary for the normal expression pattern were contained within the first 1-2kb of 5' flanking DNA (Romano *et.al.*, 1988). Further analysis of the S15 chorion protein gene from the 3rd chromosome cluster has identified three discrete regulatory elements in the 5' flanking DNA, two of which govern temporal specificity whilst the third dictates tissue specificity (Wong *et.al.*, 1985; Mariani *et.al.*, 1988).

The huge demand for chorion proteins during eggshell deposition is met by the selective amplification of the genes in the follicle cells (Spradling and Mahowald, 1980).

A putative vitelline membrane protein cDNA was isolated in a differential screen by Burke *et.al* (1987) When used as a probe in *in situ* hybridisation to ovarian tissue sections it recognised a transcript present in stage

10 follicle cells, coincident with the time of vitelline membrane protein production and the size of the protein product predicted from DNA sequencing matched a known vitelline membrane protein (Fargnoli and Waring, 1982).

1.13 Maternal effect mutations

The maternal effect mutations are mutations which when present in the maternal genome, disrupt zygotic development. The genes identified by these mutations are typically expressed in the nurse cells of the ovary and the mRNA or protein products are stored in the oocyte until required to act in early embryogenesis. Since there is virtually no transcription of the zygotic genome during the first ninety minutes of embryogenesis and very little transcription until the cessation of syncytial divisions at about 2.5 hours (Edgar and Schubiger, 1986) all the gene products required for housekeeping functions and nuclear divisions in the early hours of development must be provided maternally. The maternal genome is also required to provide regulatory gene products involved in processes of embryonic differentiation. Two areas of embryonic development in which the maternal contribution is well documented are sex-determination and the specification of the anterior-posterior and dorsal-ventral axes.

1.13.1 Sex-determination

As discussed earlier, female development and viability of XX progeny depends on activation of the key gene *Sxl* in the zygotic genome. This activation requires the products of several maternally expressed genes, two of which are *da* (Cline, 1978) and *Dk* (Steinmann-Zwicky *et.al.*, 1987). At 25°C mothers homozygous for a temperature sensitive *da* allele produce no daughters regardless of the paternal genotype at this locus (Bell, 1954). Like *da*, the dominant mutation *Dk* is a temperature-sensitive maternal effect

mutation which is lethal to almost all female progeny (Steinmann-Zwicky et.al., 1987). The lethal effects of both these mutations are rescued by the presence in the zygote of a constitutively active *Sxl* mutation (*Sxl^{M#1}*) (Cline, 1984; Steinmann-Zwicky et.al., 1987). *da* transcripts are present in early embryos (Cronmiller et.al., 1988) and the protein product has been implicated in the transcriptional activation of *Sxl*, around blastoderm (see earlier). The mechanism whereby the *Dk* product exercises its function in *Sxl* activation is undetermined.

1.13.2 Anterior-posterior axis

Genetic screens have identified several groups of maternally acting genes controlling embryonic pattern formation, on the basis of mutant phenotypes characterised by region specific pattern defects (reviewed by Nusslein-Volhard et.al., 1987). A number of these genes were subsequently cloned and from the genetic and molecular data, a clear picture has emerged of the mechanisms operating.

The initial organisation of anterior-posterior body plan is determined by three localised maternal gene products: the *bicoid* (*bcd*) product controls anterior development (Berleth et.al., 1988; Driever et.al., 1990 and the *nanos* (*nos*) and *torso* products control development of the posterior and terminal regions respectively (Sander and Lehmann, 1988; Lehmann and Nusslein-Volhard, 1991; Wang and Lehmann, personal comm. in St.Johnston et.al., 1991). A number of additional maternally acting genes are involved in the localisation of each these signals. *bcd* transcripts are synthesised in the nurse cells and upon transfer to the oocyte are trapped at the anterior pole, by components of the cytoskeleton encoded maternally by the genes *swallow* and *exuperantia* (Berleth et.al., 1988). The final stage of this localisation also requires the product of a third maternally acting gene, *staufer* (*stau*) and it has recently

been shown that *stau* protein accumulates at the anterior egg pole in the same region as *bcd* RNA and may be acting as a molecular chaperone (St.Johnston *et.al.*, 1991). Mutations in *bcd* or any of the other three genes result in anterior pattern defects (Schupbach and Wieschaus, 1986; Frohnhofer and Nusslein-Volhard, 1987; Berleth *et.al.*, 1988; St.Johnston *et.al.*, 1989). During the syncytial cleavages *bcd* RNA is translated and the protein diffuses away from the anterior pole, forming a gradient over about half the egg length (Driever and Nusslein-Volhard, 1988 a and b). The *bcd* protein has a putative DNA binding homeodomain and is associated with the syncytial nuclei (Driever and Nüsslein-Volhard, 1989). Its role in pattern formation is to control, in the anterior of the embryo, the expression of members of the zygotic gap gene class. These genes head the hierarchy of zygotic regulatory genes governing segmentation and they are responsible for the initial subdivision of the embryo, which is further elaborated by the pair rule and segment polarity genes (reviewed in Ingham, 1988). Examination of the expression pattern of the gap gene *hunchback* (*hb*) in embryos lacking the *bcd* gene product indicates that the function of *bcd* is the transcriptional activation of the zygotic *hb* gene in the anterior half of the embryo (Tautz, 1988) and the *bcd* product has been shown to bind to high and low affinity sites in the *hb* promoter (Driever and Nusslein-Volhard, 1989). The *hb* product in turn represses the transcription of the gap gene *Kruppel* (*Kr*) in the anterior region (Jäckle *et.al.*, 1986; Gaul and Jäckle, 1987).

A total of nine maternally acting genes are involved in providing the posterior signal. The large number of genes may reflect the more complex requirements of posterior determination. First, a simple trapping mechanism is impossible since the posterior pole is not adjacent to the nurse cells, so the necessary factors are required to traverse the length of the egg. Second, the posterior determination mechanism is closely integrated

with the process of pole cell determination. In addition to *nos*, the eight genes which can be mutated to yield abdominal defects are *cappucino* (*cap*), *spire* (*spire*), *staufen* (*stau*), *oskar* (*osk*), *vasa* (*vas*), *tudor* (*tud*), *valois* (*vls*) and *pumilio* (*pum*) (Schupbach and Wieschaus, 1986; Boswell and Mahowald, 1985; Lehmann and Nusslein-Volhard, 1986; Nusslein-Volhard et.al., 1987; Manseau and Schupbach, 1989). Mutants in any of the first seven of these genes produce *nos* rescuing activity in their nurse cells, as evidenced by injection of nurse cell cytoplasm into embryos of *nos* mothers, but fail to localise the *nos* product to the posterior pole. Thus the wild type products of these genes are evidently involved in the initial localisation of *nos* RNA to the posterior pole (Lehmann and Nusslein-Volhard, 1991).

It is known from transplantation experiments that the posterior determining activity is actually required in the presumptive abdominal region of the embryo rather than the posterior pole (Lehmann and Nusslein-Volhard, 1986). Embryos of *pum* mothers exhibit abdominal defects but do possess *nos* activity at the posterior pole and it appears that the *pum* product may be involved in the translocation of *nos* activity from the pole to its site of action (Lehmann and Nusslein-Volhard, 1987).

Bicaudal-D (*Bic-D*) is one of a group of four maternally acting genes that causes a double abdomen phenotype when mutated (Mohler and Wieschaus, 1986). It has been cloned and encodes a filament like protein (Suter et.al., 1989; Wharton and Struhl, 1989). The mutant phenotype is associated with the presence of *nos* activity at both poles (Lehmann and Nusslein-Volhard, 1986). A number of genes of the posterior group have been cloned, allowing a hierarchy of functions to be worked out. Antibodies raised against the protein product of the cloned *vas* gene were used to study the effects of various mutants on the distribution of *vas* protein, which is normally localised to the embryo posterior (Hay et.al, 1990; Lasko

and Ashburner, 1990).

The first genes in the pathway appear to be *capu*, *spir*, *stau* and *osk*, which are all required for initial *vas* protein localisation, whilst *tud* and *vls* are required for maintenance of this localisation later in embryogenesis. The *osk* gene has also been cloned (Kim-Ha et.al, 1991) and in situ hybridisation has established that *osk* transcripts are localised at the posterior of the oocyte and early embryo. It was further revealed that mutations in *capu* or *spir* completely abolish *osk* RNA localisation, placing them above *osk* in the hierarchy. The mutation *stau* also eliminates posterior *osk* RNA localisation, but permits the transient accumulation of *osk* RNA at the anterior of the oocyte. This anterior accumulation of *osk* RNA is also seen in wild type ovaries and is believed to represent an intermediate stage in the posterior localisation. Therefore, *stau* was placed either parallel to or upstream of *osk* but below *capu* and *spir* (Kim-Ha et.al., 1991). The *stau* gene has now also been cloned and *stau* protein, like *osk* RNA, been shown to be localised to the posterior pole in early oogenesis following a transient accumulation at the anterior of the oocyte. This led the authors to propose that association with *stau* protein is a prerequisite for the transport of *osk* RNA to the posterior pole (St.Johnston et.al, 1991). *vas*, *pum* and *nos* were placed below *osk* since none are required for *osk* RNA localisation and using an independently cloned *osk* gene as a probe Ephrussi et.al.(1991) confirmed that *tud* and *vls* are also dispensable for *osk* RNA localisation, and so below *osk*. They also looked at the effect of *Bic-D* mutations on *osk* RNA distribution and compared it with the distribution of *nos* RNA, using cloned *nos* DNA as a probe (Wang and Lehmann, 1991). Based on the striking similarity in the distributions of the two RNA species and the fact that *vas* protein localisation is unaffected in *Bic-D* embryos, they speculate that *osk* protein is directly responsible for localising *nos* RNA and that the other genes below *osk* in

the hierarchy, namely *vas*, *vls* and *tud*, are required for stable maintenance of *osk* protein and hence *nos* RNA, at the posterior pole. Kim-Ha et.al (1991) on the other hand, prefer a model in which *osk* protein directs *vas* protein to the posterior pole, rather than acting directly on *nos* RNA.

Manseau and Schupbach (1989) have proposed that *capu* and *spir* may encode components of a cytoskeletal framework involved in morphogen localisation and the *Bic-D* sequence indicates it could also contribute to such a framework. This suggests a model in which an *osk* RNA / *stau* protein complex binds the hypothetical framework at the oocyte anterior, before being transported to the posterior pole and that this localisation in turn directs the posterior localisation of *vas* protein and *nos* RNA.

Examination of the distribution of gap gene products in embryos lacking posterior group gene products indicates that the role of the posterior determinant *nos*, is to repress translation of *hb* mRNA in the embryo posterior. Eggs of wild type females contain homogeneously distributed, maternally provided *hb* mRNA (Tautz, 1987). When translation of this RNA begins at the eighth nuclear division *hb* protein forms a concentration gradient from anterior to posterior. Maternal mutations in all of the posterior group genes tested reduce or abolish this early *hb* protein gradient. The ectopic *hb* expression caused by these mutations, results in repression of transcription of the gap gene *knirps* in its posterior domain and since *knirps* is a negative regulator of *Kr*, the domain of *Kr* expression is shifted posteriorly (Gaul and Jäckle, 1987; Tautz, 1988; Wharton and Struhl, 1990).

1.13.2.1 The role of the posterior group genes in pole-plasm formation

There is evidently a shared pathway for the localisation of *nos* RNA and the formation of polar granules. Mutations in all seven of the posterior group

genes required for the initial localisation of *nos* RNA also result in failure to form polar granules or pole cells (Boswell and Mahowald, 1985; Lehmann and Nusslein-Volhard, 1986; Nusslein-Volhard et.al., 1987). Furthermore, *vas* and *stau* proteins are associated with the polar granules and *vas* protein is sequestered into forming pole cells (Hay et.al, 1988; St.Johnston et.al, 1990; Lasko and Ashburner, 1990). St.Johnston et.al.(1991) have proposed a direct role for *stau* protein in the formation and stabilisation of the polar granules and the unknown pole cell determining factors.

1.13.3 Dorsal-ventral polarity

The differentiation of the major tissue types (endoderm, ectoderm, mesoderm) is determined by the dorsal- ventral polarity of the embryo, with a cell's commitment to a particular developmental fate depending on its position on the dorsal ventral axis. As in the case of anterior posterior pattern formation, the dorsal-ventral pattern is first broadly specified by maternal gene products then further refined by the action of the zygotic genome. A total of 15 maternally active genes has been identified as having a role in this process (reviewed in Anderson, 1989). Three of these genes are evidently involved in establishing the dorsal-ventral polarity of the eggshell during oogenesis, whilst the remaining twelve are active in the translation of this initial cue into a continuous morphogenic gradient.

The earliest sign of dorsal ventral asymmetry is in the morphology of the egg, which is flat on the side corresponding to the presumptive dorsal side of the embryo and curved on the ventral side. This asymmetry reflects an unequal distribution of chorion proteins and depends on region-specific follicle cell behaviour (Mahowald and Kambyzellis, 1980). Mutations in *fs(1)K10*, cause a dorsalisation of the eggshell which is later mirrored in

the embryo, whilst *gurken* and *torpedo* mutations have a ventralising effect on the egg and embryo (Wieschaus, 1979; Schupbach, 1987). Clonal analyses have indicated that *K10* and *gurken* expression are required only in the germline but *torpedo* is expressed in the follicle cells (Wieschaus et.al., 1978; Schupbach, 1987). The *K10* locus has been cloned (Haenlin et.al., 1985) and molecular studies have revealed that the gene is transcribed in the oocyte, and encodes a putative DNA binding protein which accumulates in the oocyte nucleus (Haenlin et.al., 1987; Prost et.al., 1988). The phenotypes of double mutants infer that *torpedo* and *gurken* act downstream of *K10*. Based on these observations, plus the fact that the oocyte nucleus is located dorsally within the oocyte, a speculative model of the interaction of these loci has been proposed (Anderson, 1989). According to this model the *K10* product activates the expression of other genes in the oocyte, which in turn influence the behaviour of the adjacent (dorsal) follicle cells, in a process requiring the *gurken* product in the oocyte and the *torpedo* product in the follicle cells.

In response to the cue established by these three genes, a morphogenetic gradient is established in the embryo, by a further twelve maternally acting loci. Loss of function mutations in eleven of these genes (the dorsal group) result in completely dorsalised embryos lacking all ventral and lateral structures (Nusslein-Volhard, 1979; Anderson and Nusslein-Volhard, 1986). However, hypomorphic mutations, which have been isolated for most of these loci (K.V.Anderson and S.Wasserman, unpublished) only eliminate the most dorsal structures. These results infer a gradient in the requirement for these gene products, with the highest concentrations required for the differentiation of ventral structures. Temperature shift experiments and patterns of epistasis observed in double mutant combinations, have been informative with the respect to the order of function of the dorsal group genes. Based on the temperature sensitive periods of mutant alleles, five

genes; *gastrulation defective*, *nudel*, *pipe*, *spatzle* and *windbeutel*, seem to be required during late oogenesis, a conclusion reinforced by the fact that these mutants are not rescued by injection of wild type cytoplasm after fertilisation. On the same criteria, *tube*, *snake*, *easter*, *pelle*, *Toll* and *dorsal* all act during the syncytial blastoderm stage (Anderson, 1989). Loss of function alleles of *cactus* and dominant gain of function alleles of *Toll* (*Toll^D*) have ventralising effects on embryos. By examining the patterns of epistasis in combinations of these alleles with dorsalising loss of function alleles of the other genes, it has been possible to derive the sequence in which these genes function (Anderson et.al., 1985 and Anderson, unpublished in Anderson, 1989). For example, *dorsal* is deduced to act downstream of *Toll* because the dorsalising phenotype of recessive *dorsal* mutations is epistatic to the *Toll^D* phenotype and *Toll* is placed downstream of *gastrulation defective*, *snake* and *easter* because the *Toll^D* phenotype is epistatic to their mutant phenotypes. Similarly, since *cactus* is epistatic to all the other genes except *dorsal*, *dorsal* appears to act at the end of the pathway, preceded by *cactus*. Dorsalized phenotypes resulting from mutations in the seven genes identified as acting during blastoderm, can be rescued by injection of wild type cytoplasm into the affected embryos (Anderson and Nüsslein-Volhard, 1986; Anderson et.al., 1985). However, in all mutants except *Toll*, the ventralising effect of wild type cytoplasm is independent of the site of injection and the ventral most structures always differentiate at the wild type position i.e. from the cells lying on the curved presumptive-ventral side of the egg (Anderson and Nüsslein-Volhard, 1986). By inference, the wild type products of the genes concerned must have a role in interpreting a pre-existing gradient of information rather than establishing that gradient. In contrast, in *Toll* embryos ventral-most structures are always formed at the site of injection, even if that is the presumptive dorsal side of the egg (Anderson

et.al., 1985). Thus *Toll* seemed a candidate for encoding the morphogen underlying dorsal-ventral polarity, until the gene was cloned and it was determined that *Toll* encodes a transmembrane protein which is evenly distributed throughout the egg membrane (Gerttula *et.al.*, 1988 and Hashimoto *et.al.*, 1988). To reconcile these contradictory results, it must be assumed that although *Toll* protein is present throughout the egg, its activity is spatially regulated in wild type embryos by an asymmetric cue. Recently, it has been shown that the expression of three of the dorsal group genes, *nudel*, *pipe* and *windbeutel* is required in the follicle cells (Stein *et.al.*, 1991) and it has been proposed (Nüsslein-Volhard, 1991) that their products, perhaps in association with the products of *K10*, *gurken* and *torpedo* are responsible for the provision of a ventrally localised ligand which activates the *Toll* receptor in the presumptive ventral region of the embryo. Another two dorsal group genes, *snake* and *easter* have been shown to encode secreted serine proteases (DeLotto and Spierer, 1986 ; Chasan and Anderson, 1989) which are candidates for effecting a timed release of the putative ligand from the vitelline membrane, in the early stages of embryogenesis. Assuming this speculative regulatory system is correct, it can evidently be overridden by artificially high local concentrations of *Toll* protein, as shown by the cytoplasmic injection experiments. These experiments indicate that *Toll* protein may have some autocatalytic function, a conclusion previously inferred from the interactions between certain *Toll* alleles (Anderson *et.al.*, 1985). A model has been proposed which accommodates the data, in which following the initial ventral activation of the *Toll* receptor in wild type embryos, the resulting signal is maintained and amplified in an autocatalytic loop involving the *Toll* product itself, with the extracellular domain acting as both ligand and receptor. According to this hypothesis, in the injection experiments there is self activation of the *Toll* product at the site of injection,

bypassing the normal positional cue (Anderson, 1989).

The product of *dorsal*, which is evidently at the end of the pathway is probably a transcription factor since it is a nuclear protein with similarity to the avian oncogene *rel* (Steward, 1987; Steward *et.al.*, 1988) and the transcription factor NF κ B (Kieran *et.al.*, 1990; Ghosh *et.al.*, 1990). *dorsal* mRNA (Steward *et.al.*, 1985) and protein (Roth *et.al.*, 1989; Rushlow *et.al.*, 1989; Steward, 1989) is uniformly distributed in the newly layed egg. However, by the syncytial blastoderm stage, *dorsal* protein is concentrated within nuclei at the ventral side of the embryo but remains in the cytoplasm at the dorsal side (Roth *et.al.*, 1989; Rushlow *et.al.*, 1989; Steward, 1989). In dorsalised mutant embryos there is a failure in the nuclear uptake of *dorsal* protein at the ventral side, whilst ventralised *cactus* mutants exhibit *dorsal* protein uptake by both ventral and dorsal nuclei (Roth *et.al.*, 1989). Therefore it appears that *cactus* functions as an inhibitor of *dorsal* protein uptake and the genes of the dorsal group stimulate the process. In accordance with this information, the role of the activated *Toll* receptor could be to somehow override the inhibitory effect of the *cactus* protein in the ventral region.

The generation of dorsal-ventral pattern in response to the graded information provided by the maternally acting genes is mediated by a number of zygotically expressed regulatory loci including *twist*, *snail*, *zerknüllt* and *Notch* (reviewed in Anderson, 1987) and in the cases of *twist* and *zerknüllt* there is evidence that *dorsal* may be directly involved in regulation of their expression (Thisse *et.al.*, 1986; St.Johnston and Gelbart, 1987).

1.13.4 Determination of terminal pattern elements

The acron and telson are unsegmented regions at the anterior and posterior ends of the embryo respectively and their formation depends on a third group of maternally

acting genes, namely *torso*, *torsolike*, *trunk*, *fs(1) Nasrat*, *l(1) polehole* and *fs(1) polehole*. Females homozygous for loss of function alleles of any of these terminal group genes produce embryos lacking both acron and telson (Reviewed by Lehmann, 1988). Temperature shift experiments with temperature sensitive alleles have indicated that *torso* and *trunk* are active exclusively after fertilisation (Schüpbach and Wieschaus, 1986) whereas *torsolike* functions during oogenesis (Frohnhofer, 1987 in Anderson, 1989) and the fact that null alleles of *fs(1) Nasrat* and *fs(1) polehole* cause collapsed eggs (Perrimon et.al., 1986 and Degelman et.al., 1986) suggests that these genes also have an early function.

The existence of dominant gain of function alleles of *torso*, which result in expansion of the terminal regions, infer a key role for this gene in directing terminal region formation (Klingler et.al., 1988). A *torso* product is evidently required in the terminal cytoplasm since the phenotype of *torso* loss of function mutations can be rescued by the injection of wild type cytoplasm at the terminal regions of the embryo, although the rescuing activity is uniformly distributed in the wild type embryo. These observations led to the hypothesis that *torso* function depends on localised activation in the terminal regions of a uniformly distributed inactive precursor (Klingler et.al., 1988). More recently, the *torso* gene has been cloned and it has been confirmed that the protein is evenly distributed (Sprenger et.al., 1990). Moreover, the predicted protein product resembles a receptor tyrosine kinase, implying that a localised ligand could be responsible for the activation in the terminal regions. Patterns of epistasis observed in double mutant combinations indicate that *torsolike*, *trunk*, *fs(1) Nasrat* and *fs(1) polehole* all act upstream of *torso* (Stevens et.al., 1990) while *l(1) polehole* acts downstream (Ambrosia et.al., 1989). Any of the first four genes could therefore encode the postulated ligand but recent data points to

torsolike as a likely candidate. The generation of clones of follicle cells lacking *torsolike* function, by X-ray induced somatic recombination, revealed that the *torsolike* product is only required in small populations of follicle cells at the posterior pole, for telson formation (Stevens et.al., 1990). The *torso* product presumably acts by regulating the activity of other proteins. Genetic evidence suggests that the zygotically active gene *tailless* could encode one of its targets since the *tailless* mutant phenotype is similar to the loss of function phenotype of the terminal group genes and is epistatic to the phenotype of the dominant *torso* alleles (Strecker et.al., 1988).

Although development of the telson appears independent of the posterior group genes (Nüsslein-Volhard et.al., 1987) acron formation is dependent on some interaction between anterior group and terminal genes as evidenced by the fact that in the absence of maternal *bicoid* activity embryos develop a telson at both poles (Frohnhofer and Nüsslein-Volhard, 1986).

1.14 Enhancer-trap gene detection system

In order to further advance our understanding of the molecular basis of sexual differentiation it will be necessary to clone and analyse the expression of as many of the genes involved as possible. Traditionally, genes of interest have been identified on the basis of a mutant phenotype which indicates a function in the process under investigation. The cloning of these genes typically involves a long and laborious series of steps. In the first place, the physical location of the gene is determined approximately by recombination mapping, then it is more precisely localised by deficiency mapping and cytological examination of chromosomal rearrangements which disrupt the gene function. Finally, a chromosome walk is initiated from the nearest cloned region until the desired transcription unit is located. In addition to the protracted isolation procedure, conventional mutagenesis suffers from the limitation of selectivity for genes with an informative mutant phenotype. A proportion of genes, e.g. pleiotropic genes or those with an early lethal mutant phenotype or a redundant function which can be complemented by another locus, will not be detectable by this method. Many genes involved in sexual differentiation may escape detection. As discussed previously, genes under the regulation of the sex-determination hierarchy will not necessarily have sex-specific mutant phenotypes and genes controlling the formation of the reproductive systems are unlikely to be detected by screening for an adult viable, sterile phenotype since many will have other vital functions in development. In this connection, it should be noted that most of the genes involved in oogenesis also appear to have functions in other tissues or at other developmental stages (Perrimon *et.al.*, 1986). The recently developed system of P-element mediated enhancer detection (O'Kane and Gehring, 1987) offers an alternative approach to gene detection and isolation which promises to overcome some of

these problems. The technique involves random P-element mediated integration into the *Drosophila* genome of a reporter gene, which displays a pattern of expression dictated by adjacent endogenous *cis*-acting enhancer elements. The pattern of reporter gene expression therefore mirrors the expression pattern of a gene nearby. A single transformed line can be used to generate multiple lines with independent insertion sites, by crossing the parent transformant to a stock carrying the Δ 2-3 "jumpstarter" element at chromosomal position 99B (Robertson *et.al.*, 1988). The reporter gene used consists of *E.coli lac-Z* sequences fused in-frame to the second exon of the P-element transposase gene and is under the P-element promoter, which confers a weak background of constitutive β -galactosidase activity, against which the specific patterns of expression can be observed (Fig 1.3). This system, which allows genes of interest to be detected on the basis of their wild-type expression patterns, rather than relying on a mutant phenotype, should give access to genes undetectable by mutagenesis, which can then be readily cloned. With the first generation enhancer detection construct (O'Kane and Gehring, 1987) (Fig 1.3) the cloning procedure involved the construction of a genomic library, as in conventional transposon tagging (Searles *et.al.*, 1983), but this was simplified with the introduction of a modified version of the vector, which includes the *E.coli* origin of replication and β -lactamase gene. This facilitates plasmid rescue of the DNA flanking the insertion (Bellen *et.al.*, 1989; Bier *et.al.*, 1989).

Results from a number of studies have demonstrated the validity of the enhancer detection approach. Bier *et.al* (1989) examined the embryonic stages of over 3700 independent transformed lines and found that around 60% exhibited a spatially restricted pattern of β -galactosidase activity, indicating that the reporter gene was under the influence of endogenous regulatory elements. Bellen *et.al* (1989) in a survey of embryonic staining in over 500 lines

Figure 1.3

Structure of the enhancer trap transposon p[*lac Z*, *ry*⁺]

(modified from O'Kane and Gehring, 1987)

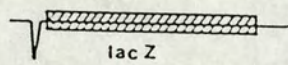
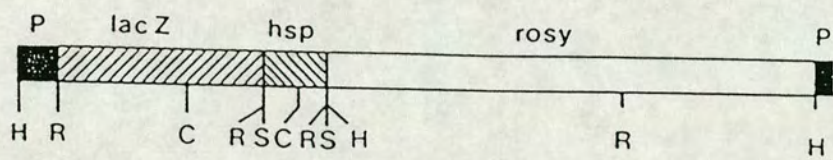
H = Hind III


R = Eco RI


C = Sac I


S = Sal I


The 3' end of the *E. coli lac Z* gene is fused in frame to the second exon of the P transposase gene and transcription of the hybrid gene is from the weak P-element promoter situated at the extreme left in the diagram. The construct includes the stop codon and polyadenylation site of the *Drosophila hsp 70* gene, P-element terminal repeats and the *Drosophila rosy* gene as a visible marker. In lines transformed with this construct, digestion of genomic DNA with Sal I yields fragments comprising the 5' end of the construct, including the *lac Z* sequences, plus a variable amount of flanking genomic DNA.



 P-Element (Carnegie 20)

 lac Z

 hsp 70

 rosy⁺

reported a similar proportion. Two lines of evidence suggest that in a significant proportion of cases the observed staining patterns reflect the expression of genes in the vicinity (Reviewed in Wilson et.al., 1990).

1) The chromosomal positions of the insertions in around 200 lines, the majority of which exhibit spatially restricted staining patterns, have been determined by in situ hybridisation to polytene chromosomes. In at least 15 cases the staining pattern corresponds with the expression of a previously characterised gene which maps to the same location, in spite of the fact that the expression patterns are known for only 1% of *Drosophila* genes. Also, Wilson et.al (1989) have described two lines with staining patterns faithfully reproducing the expression of the collagen type IV and fasciclin III genes, which were determined by molecular analysis to lie immediately downstream of the enhancer-trap insertions.

2) In a survey of 23 transformed lines, Wilson et.al (1989) cloned fragments of genomic DNA immediately 3' of the insertion by plasmid rescue and these fragments (or complementary cDNA clones) were labelled and used to probe tissue sections. Of the seven probes which detected signals above background, four detected a pattern of transcripts matching the β -galactosidase staining pattern of the corresponding transformant. Considering the fact that only a small region of genomic DNA at one side of the transposon was examined these results almost certainly provide an underestimate of the proportion of staining patterns reflecting the expression of a nearby gene.

The validity of this approach is also supported by the enormous variety of staining patterns reported in the enhancer trap screens undertaken to date, which indicates that an extensive range of genes should be accessible. In embryos, staining may be confined to particular tissues, organs or cell types. Complex tissues composed of many different cell types, such as the nervous system display a particularly rich variety of staining patterns. In a number

of cases the staining patterns have identified previously unrecognised cell types (Bier *et.al.*, 1989; Bellen *et.al.*, 1989). Several examples of pole cell or gonad specific staining have been reported. A number of strains will provide valuable tissue or cell type specific markers for developmental studies. For example transformed lines are available which allow independent labelling of virtually every cell type in the peripheral nervous system. A number of region-specific patterns, some segmentally repeated, are also observed, which may identify genes responding to positional information. Bier *et.al* (1989) determined that the insertions in 12% of the lines which expressed β -galactosidase were associated with recessive lethal mutations and noted a correspondence between the tissues expressing β -galactosidase and the tissues affected by the respective mutations.

A similar complexity of staining patterns is seen in later developmental stages, offering the potential for analysing the development of complex structures such as the brain, imaginal discs and reproductive organs (Gibson and Gehring, 1990; Bellen *et.al.*, 1990). Grossniklaus *et.al* (1989) have applied the enhancer detection system to the study of oogenesis, examining the ovarian staining patterns in nearly 600 lines. About 47% of these exhibit some ovarian staining, presumably reflecting the high proportion of the genome required to execute such a complicated process as oogenesis. However, only two of these lines stain exclusively in ovaries, in keeping with the genetic evidence of Perrimon (1986) that most genes involved in oogenesis also have other functions. Of the 19% which stained in germline cells most were stained only in nurse cells while only one line stained solely in the oocyte. A few lines showed different levels of staining in different nurse cells suggesting that all nurse cells are not functionally equivalent. A range of temporal regulation was observed which suggested that only 15% of genes with a function in the nurse cells are active before stage 6 of

oogenesis and of the 85% active at later stages, most begin expression around stage 9 and 10.

10% of lines expressed β -galactosidase exclusively in the follicle cells, mostly in sub-populations which have been previously identified morphologically. However, in a few cases previously unrecognised subgroups of cells were stained. For example, one line stained in a belt of follicle cells around the middle of the oocyte. Anterior-posterior and dorsal-ventral gradients of staining were also observed in follicle cells, possibly reflecting the expression of genes with roles in embryonic pattern formation. Based on the observed staining patterns, the time course of gene activity in the follicle cells is similar to that in the nurse cells, with only a few genes active in the early stages of oogenesis and a burst of new gene activity at stages 9 and 10.

In a screen of transformants carrying insertions on the X or second chromosomes, eight insertions were found to be associated with recessive mutations causing sterility or reduced fertility. For seven of these, it was ascertained that the mutation was caused by the insertion, because a high frequency of wild type revertants could be obtained by remobilisation of the transposon. Moreover all seven lines were found to express B-galactosidase in their ovaries, specifically in nurse cells, suggesting that the staining patterns did reflect the expression of the disrupted gene.

On the above evidence the enhancer trap system is confirmed as a potentially powerful research tool. The range of staining patterns generated will provide valuable cell type or position dependent markers for developmental studies and in a proportion of cases will identify interesting genes which can then be readily cloned. Furthermore, the enhancer trap vector can be used for insertional mutagenesis and the information gained from a recessive mutant phenotype will be supplemented by knowledge of the wild type gene expression pattern, deduced

from the B-galactosidase activity in heterozygotes. The vector can also facilitate the mutational analysis of a nearby gene of interest by the generation of deletions or other rearrangements through imprecise excision.

1.15 Cloning of Drosophila IMP dehydrogenase

This thesis describes the investigation of an enhancer-trap line exhibiting testis-specific β -galactosidase staining. Clones of the genomic region around the insertion site were isolated from a genomic library and were found to contain part of a transcription unit encoding multiple overlapping RNA's. One of the transcripts is testis-specific and another is highly enriched in ovaries and early embryos. The predicted protein sequence of a cDNA clone representing the ovary-enriched RNA species, contained a region of very high similarity to the published sequences of human and mouse inosine-5'-monophosphate dehydrogenase (IMPD), which is the rate limiting enzyme in *de novo* GTP synthesis and has been implicated in human cancer. Two other cDNA's, one encoding the full length *Drosophila* IMPD and the other putatively representing the testis specific transcript, were also isolated and sequenced. The *Drosophila* IMPD gene was further characterised by Northern and in situ hybridisation techniques and the significance of the observed expression pattern is discussed in relation to the known functions of the enzyme.

CHAPTER TWO

MATERIALS AND METHODS

2.1 Solutions

Chemicals were obtained from Sigma, BDH and Aldrich.

Enzymes (Restriction Endonucleases, T4 DNA Ligase, Klenow fragment of DNA Polymerase, DNase I, RNase A, Proteinase K) were obtained from BRL, Pharmacia and NBL.

Radioisotopes were obtained from Amersham.

Hybond-N nylon membranes were obtained from Amersham.

Standard solutions were made using sterile distilled water in sterile, baked glassware and were generally made sterile by autoclaving (15psi/15 min). Solutions not detailed in the text are described below.

TE	10mM Tris-Cl pH 7.4-8.0, 1mM EDTA
----	-----------------------------------

0.5M EDTA	0.5M Diaminoethanetetra - acetic acid, pH 8.0
-----------	---

Boehringer Mannheim Incubation Buffers for Restriction Enzymes :

Final concentrations in mm/L (i.e. 1:10 diluted buffer).

A	:	33mM Tris HCl, 10mM Mg-Acetate, 66mM K-Acetate, 0.5mM Dithiothreitol (DTT), pH 7.5
---	---	--

B	:	10mM Tris HCl, 5mm MgCl ₂ , 100mM NaCl, 1mM 2-Mercaptoethanol, pH 7.5
---	---	--

L	:	10mM Tris HCl, 5mM MgCl ₂ , 1mM Dithioerythritol (DTE), pH 7.5 (contd)
---	---	---

M	:	10mM Tris HCl, 10mM MgCl ₂ , 50mM NaCl, 1mM DTE, pH 7.5
---	---	--

H	:	50mM Tris HCl, 10mM MgCl ₂ , 100mM NaCl, pH 7.5
---	---	--

10X TBE gel buffer	:	0.89M Tris-Borate, 0.89M Boric Acid, 10mM EDTA
--------------------	---	--

10X DNA and RNA :
 gel loading buffer 100mM EDTA pH 8.0, 0.1% (w/v)
 Bromophenol Blue, 20% (w/v)
 Ficoll (Type 400)

10X MOPS gel buffer : 0.2M Na-MOPS pH 7.0, 50mM Na-Acetate, 10mM EDTA

RNA Formaldehyde :
 Sample Buffer (FSB) 50% (v/v) formamide, 25% (v/v) formaldehyde (at 4% w/v), 25% (v/v) 10X MOPS buffer

RNA extraction buffer : 100mM Tris HCl pH 7.5, 10mM EDTA, 150mM Lithium Chloride, 1% (w/v) SDS

OLB (oligo labelling buffer)
 Solution 0 : 0.125M MgCl₂, 1.25M Tris HCl pH8
 Solution A : 0.95ml solution 0, 18ml 2-Mercaptoethanol, 25μl 20mM dATP, 25μl 20mM dTTP, 25μl 20mM dGTP
 Solution B : 2M Hepes pH 6.6
 Solution C : Hexadeoxyribonucleotides (Pharmacia) suspended in TE buffer at 90 OD units/ml
 OLB is made by mixing Solution A, B and C in the ratio 2:5:3 (and is stored at -20°C)

20X SSC : 3M NaCl, 0.3M Na- Citrate

20X SSPE : 3.6M NaCl, 20mM NaHO₄ pH 7.4, 20mM EDTA pH 7.4

Salmon Sperm DNA : 10mg/ml stock which has been sonicated, phenol extracted and ethanol precipitated. Purchased from SIGMA.

50X Denhardt's Solution : 1% (w/v) Ficoll, 1% (w/v) polyvinyl pyrrolidine, 0.2% (w/v) BSA

Ringer's Solution	:	3.2g NaCl, 3.0g KCl, 1.8g MgSO ₄ , 0.69g CaCl ₂ .2H ₂ O, 1.79g Tricine, 3.6g glucose, 17.1g sucrose. Made up to 1 litre with dH ₂ O, pH adjusted to 6.95, filter sterilised and stored at 4°C
10X PBS	:	0.1M Sodium Phosphate pH 7.5, 1.3M NaCl
TM buffer	:	10mM Tris HCl pH 8.0, 10mM MgCl ₂
TE buffer	:	10mM Tris HCl pH 7.4-8.0, 1mM EDTA
NET	:	150mM NaCl, 0.1mM EDTA, 20mM Tris HCl pH 8.0
HNET	:	1.0M NaCl, 0.1mM EDTA, 20mM Tris HCl pH8.0

Table 2.1: List of *Drosophila melanogaster* Strains and Balancer chromosomes used and Their Relevant Features

Stock	Relevant Features	Reference/Source
Oregon R (OrR)	A wild-type strain	Lindsley and Grell (1968)
w; Sb e Δ 2-3 / TM6	stable source of P transposase for mobilisation of p[lac-Z ry ⁺].	J.Merriam
T (2,3) CyO TM6 / pr cn; mwh ry e	stock used for generation of homozygous p[lac-Z, ry ⁺] transposants.	J.Merriam
Balancer Chromosomes		
TM6	In (3LR) TM6, Hn ^p ss ^{P88} bx ^{34e} Ubx ^{P15} e	
T (2,3) CyO TM6	reciprocal translocation between balancer chromosomes CyO and TM6. Suppresses crossing over in chromosomes 2 and 3.	

2.2 Maintenance of *Drosophila* Stocks and Strains

Drosophila stocks were maintained in vials or bottles at 18°C or 25°C on cornmeal food which consisted of cornflour 250g, sugar 500g, yeast pellets 175g and agar 100g made up to 10 litres with dH₂O. The food was then boiled, cooled to approximately 40°C and poured into bottles or vials and allowed to set. A fungicide, Nipagen was added to a final concentration of 4.5µg/L and on occasion antibiotics such as Gentamycin (to 40µg/ml) were added. Strips of filter paper that had been soaked previously in 3% (v/v) Benzyl Benzoate in ethanol and

allowed to air dry, were placed on top of the food if mite infections occurred.

Table 2.2: Stages of Drosophila melanogaster development at 25°C (Reference: Bownes and Dale, 1982)

Hours	Days	Developmental Stage
0	0	Fertilization and fusion of pronuclei
1.5		Preblastoderm stage. Migration of cleavage nuclei and pole cell formation.
3		Blastoderm stage. Migrated nuclei from cells in the previously syncytial blastoderm.
3.5		Gastrulation begins
6-8		Segmentation visible
18		Larval differentiation nearly complete
24	1	Hatching from egg. Onset of first larval instar.
48	2	First moult. Second larval instar begins.
72	3	Second moult. Third larval instar begins.
120	5	Puparium formation with white puparium
122		Puparium darkens
124		Prepupal moult
132		Pupation. Eversion of imaginal discs
216-240	9-10	Emergence of adult from pupal case

2.3 Drosophila melanogaster Life Cycle - Collection of Developmental Stages

2.3.1 Stages of Drosophila melanogaster Development at 25°C

Table 2.2 summarises the stages of *Drosophila melanogaster* development at 25°C for the strain used, Oregon R.

2.3.2 Collection of Eggs (0-6 hour embryos)

Eggs were collected on 3% (w/v) agar plates with a smear of yeast paste (dried yeast dissolved in 20% (w/v) glucose) on the surface. The eggs were harvested by washing the collection plate with water and using a paintbrush to transfer the eggs to a sieve where any traces of food were removed by rinsing with water. The eggs (which represented a collection of 0-6 hour embryos) were frozen in liquid nitrogen and stored at -70°C .

2.3.3 Collection of Late 3rd Instar Larvae and Pupae

Egg laying flies were placed in fresh food bottles for 4 hours then removed. The bottle was then incubated for 96 hours and 120-240 hours. Late 3rd instar larvae were picked with a paintbrush as they crawled up the sides of the bottles after an incubation period of 96 hours, placed in Eppendorf tubes and frozen with liquid nitrogen, and stored at -70°C .

Pupae were collected by picking them off the inside of the food bottles after 120-140 hours incubation, placed in Eppendorf tubes and frozen in liquid nitrogen, and stored at -70°C .

2.3.4 Collection of Sexed Adults

Adult flies were collected ranging in age from newly eclosed to 10 day old flies. The flies were anaesthetised with di-ethyl ether, sexed, placed in Eppendorf tubes, frozen in liquid nitrogen and stored at -70°C .

Table 2.3: List of Microbial Strains, Vectors, Libraries, Plasmids and Recombinant bacteriophage Used.

	Relevant Genotype	Comments	Reference
TG1	(lac,pro), supE thi,.hsds/F' tra D36, proA ⁺ B ⁺ , lacI ^s , ZΔ M15	Used as host for M13 bacteriophage	Amersham International (unpublished)
NM514	hfl, hsdR	selection for cI ⁻ recombinants in λgt10 male cDNA library	Murray (1983)
XL-1 blue	Rec A1, end A1, gyr A96, thi, hsdR17, supE44, rel A1, λ- Δ(lac), {F' proAB, lacI ^s , ZΔM15, Tn10(tet ^R)}	Used to screen λZAP cDNA library and propagate pBluescript clones	Bullock <i>et.al.</i> , (1987)
Y1090	hsdR supF Δlac Δlon pMC9	contains pMC9 which is amp ^r and produces high levels of lac repressor. Deficient in lon protease. Used with λ gt11	Young and Davis (1983)
C600	F ⁻ thi-1 leub6 lacY ⁺ tonA21 supE44	used for growth of λ	Appleyard (1956)
Q359	hsdR hsdM ⁺ supF φ80 P2	used for detection of Spi ⁻ recombinants of λ.	Karn <i>et.al.</i> , (1980). (contd)

Plasmids	Relevant Genotype	Comments	Reference
pGEM1/2	amp ^r	pGemini vector used for subcloning cDNA's and generating RNA probes	Promega Biotech. Melton (1984)
Gem α -tubulin	amp ^r	pGemini-2 plasmid containing a 1.7 kb α -tubulin-1 DNA fragment. This plasmid was constructed in this laboratory by M.Blair from a plasmid gifted by Kalfayan and Wensink	Kalfayan and Wensink (1982)
pUC 18/19	amp ^r	used for subcloning cDNA's and double stranded sequencing-offers colour selection	Messing (1983)
pBluescript	amp ^r	used for generating single stranded sequencing templates	

pMC1871	Tet ^r	3.1 kb fragment comprising entire coding sequence of <i>E.coli lac Z</i> gene minus the first seven codons cloned into the Pst 1 site of pBR322	Casabadan <i>et.al.</i> , (1983)
---------	------------------	--	--

Bacteriophage	Comments	Reference
M13mp18/19	vectors used for subcloning and generation of single stranded sequencing templates	Yannish-Perron (1983)
EMBL4	λ replacement vector used for construction of <i>Drosophila</i> genomic libraries	Frischauf <i>et.al.</i> , (1983)

2.3.5 Libraries

2.3.5.1 Canton S genomic library

This random genomic library constructed from the strain Canton S was a gift from the laboratory of D.M Glover (University of Dundee). It was constructed by partial digestion of genomic DNA with Sau 3A1 and cloning into the Bam H1 site of the λ replacement vector λ DASH II (Stratagene). The donor DNA was first size selected to ensure inserts in the range 15-20 kb. Further details of the construction procedure were unavailable. The λ DASH II vector possesses T7 and SP6 promoters flanking its multiple cloning site which permit the transcription of cloned sequences (section 2.5.8.5). The library was plated on *E.coli* strain Q359 which carries a P2 lysogen to select

against non-recombinant phage.

2.3.5.2 Testis cDNA library

This library was constructed by Stratagene in the vector λ ZAPII using both oligo dT and random primed first strand synthesis on poly (A)⁺ RNA extracted from adult *Drosophila* testes. The λ ZAP II vector permits the excision of a pBluescript plasmid containing the cDNA insert by co-infection with a helper phage (section 2.5.5.5). The library was plated on the recommended XL-1 Blue cell strain (Table 2.3).

2.3.5.3 Adult male cDNA library

This library was a gift from Thomas Kornberg (UCLA). It was prepared from poly(A)⁺ RNA extracted from whole adult males using oligo dT primed first strand synthesis and the insertion vector λ gt10. The library was plated on the hfl *E.coli* strain, NM514 to eliminate non-recombinant phage.

2.3.5.4 Ovary cDNA library

poly(A)⁺ RNA was extracted from whole ovaries of Canton S adult females. First strand synthesis was by oligo dT priming and the cDNA was cloned into the vector λ gt11. The library was screened with host strain Y1090 on L-amp plates. The library was constructed and donated by P.Sullivan and L.Kalfayan, University of North Carolina.

2.4 Microbial Strains and Media

2.4.1 Microbial Strains

All bacterial strains, plasmid vectors, bacteriophage, plasmids and libraries used are listed in table 2.3.

Bacterial stocks were maintained on appropriate plates at 4°C and also as stabs at room temperature in the dark.

Lambda bacteriophage were stored at 4°C in phage

buffer with a few drops of chloroform added.

Plasmid and Lambda DNA samples were stored suspended in TE buffer and kept at -20°C.

2.4.2 Media

All media used (listed in table 2.4) were sterilised by autoclaving. Antibiotics, vitamins and sugars were made up in dH₂O and filter sterilised. Where appropriate, antibiotics were added to plates and media, e.g. ampicillin to a final concentration of 50µg/ml.

Table 2.4 Media

L-broth	10g Difco Bacto Tryptone, 5g Difco Bacto Yeast Extract, 5g NaCl per litre, pH 7.2
L agar	16g Difco Bacto Tryptone, 10g Difco Bacto Yeast Extract, 5g NaCl per litre
BBL agar	10g Baltimore Biological Laboratories trypticase, 5g NaCl, 10g Difco agar per litre
BBL top agar	As BBL agar except, only 6.5g Difco agar per litre
Top agarose	0.7% (v/v) Pharmacia agarose in L-broth
Phage buffer	3g KH ₂ PO ₄ , 7g Na ₂ H PO ₄ , 5g NaCl, 10ml 0.1M MgSO ₄ , 10ml 0.1M CaCl ₂ , 1ml 1%(w/v) gelatin per litre

2.5 Methods

2.5.1 General Methods

2.5.1.1 Phenol Extraction

Phenol was redistilled and saturated with water (purchased from Raithburn Chemicals). It was then equilibrated with TE buffer and 8-hydroxyquinoline was added to 0.1% (w/v). Solutions of nucleic acid were deproteinised by shaking, or vortexing, with an equal volume of phenol or phenol/chloroform (1:1). The phases were then separated under centrifugation in a Sorval RC-5B centrifuge (5 Krpm, 10 minutes) or a microfuge (12 Krpm, 5 minutes) and the aqueous layer carefully removed to a fresh tube.

2.5.1.2 Ether Extraction

Diethyl ether was equilibrated with TE buffer. Solutions were shaken, or vortexed, with an equal volume of di-ethyl ether to remove any remaining phenol after deproteinisation. The phases were allowed to settle before the ether phase was removed. Any remaining ether was removed by heating the sample to 65°C for 2 minutes.

2.5.1.3 Precipitation of Nucleic Acids

The solution was adjusted to 0.3M Sodium Acetate pH 5.5 then 2.5 volumes of ethanol or 0.6 volumes of propan-2-ol were added.

The samples were incubated at -20°C for a minimum of 60 minutes. Small quantities of nucleic acids (if less than 5µg) were precipitated at -20°C overnight. The precipitate was recovered by centrifugation in a Sorval centrifuge (10 Krpm, 20 minutes, 4°C) or a microfuge (12 Krpm, 10 minutes). The pellet was washed in 70% (w/v) ethanol, dried under vacuum and re-dissolved in TE buffer or sterile, distilled water.

2.5.1.4 Estimation of DNA and RNA Yields

The optical density (OD) at 260nm was used to quantify nucleic acids. An OD_{260} of 1.0 is 50 μ g/ml for DNA and 40 μ g/ml for RNA. Nucleic acid was free of contaminating protein if the OD_{260}/OD_{280} was 1.8-2.0. Estimates of nucleic acid concentrations were also obtained by ethidium bromide staining of agarose gels with known quantities of material.

2.5.1.5 Restriction Endonuclease Digestion of DNA

In general, DNA was digested with a 2-4 fold excess of the enzyme (2 - 4 units of enzyme/ μ g of DNA) for a minimum of 60 minutes in reaction buffers purchased from Boehringer Mannheim (Section 2.1). Reaction conditions recommended by the manufacturers were used.

2.5.1.6 Extraction of DNA from Agarose Gels

a) Low melting point agarose gels: DNA fragments were fractionated on low melting point agarose gels and the desired fragment was cut from the gel. An equal volume of TE buffer was added to the sample and heated to 65°C to melt the agarose. The sample was then phenol extracted (2.5.1.1). The aqueous layer containing the DNA fragment was then ethanol precipitated (2.5.1.3).

b) DEAE Paper: DNA fragments were separated on agarose gels (2.5.6.1). The DEAE membrane (Schleicher and Schnell) was cut to the same width as the gel slot, soaked in NET buffer (150mM NaCl, 0.1mM EDTA, 20mM Tris pH 8.0) and placed into a cut in the gel, just ahead of the DNA fragment of interest. The fragment was electrophoresed onto the DEAE membrane at 100 volts for 10 minutes. The DEAE membrane was then removed from the gel, and placed in an Eppendorf tube.

The DNA was eluted from the DEAE membrane by heating at 65°C for 10 minutes in 300 μ l of HNET buffer (1M NaCl, 0.1mM EDTA, 20mM Tris pH 8.0). The eluate was removed to a fresh tube, phenol:chloroform extracted

(2.5.1.1) followed by an ether extraction (2.5.1.2). The DNA fragment was then ethanol precipitated (2.5.1.3).

2.5.1.7 Ligation of DNA Molecules

The vector and DNA insert were cut to completion with the appropriate endonucleases. Vector molecules were prevented from self ligation by removing the terminal 5'-phosphate groups by the addition of 0.01 units of calf intestinal alkaline phosphatase 5-15 minutes before digestion time was complete. Prior to ligation, the vector and insert DNA solutions were usually deproteinised by phenol extraction (2.5.1.1) and ethanol precipitated (2.4.1.3). Typically, between 10-20ng insert DNA fragment was ligated in a reaction with vector to insert DNA concentration in a 3:1 molar ratio. Ligations were carried out in a 10 μ l reaction volume containing 20mM Tris HCl pH 7.6, 5mM MgCl₂ and 5mM DTT, supplemented with 1mM dATP, and incubated overnight at 4°C. Ligations required 0.01 unit of T4 DNA for the ligation of cohesive ends. The ligation products were then transformed into *E.coli* (2.5.1.8).

2.5.1.8 Transformation into *E.coli*

L-broth (100ml) was inoculated with 1ml of an overnight culture of the relevant bacterial strain and grown at 37°C with shaking until the OD₆₅₀ reached 0.45-0.55. The culture was chilled on ice for 10 minutes. Cells were made competent and transformed with plasmid DNA by using either of the following methods:

- a) CaCl₂ Method: The cells were pelleted by centrifugation (2 Krpm, 10 minutes, 4°C) and resuspended in 50ml of ice-cold, sterile 50mM CaCl₂. The suspension was left on ice for 15 minutes and the cells were re-pelleted and resuspended in 2.5ml of ice-cold 50mM CaCl₂. One 0.1ml aliquot of this suspension was removed into a 5ml glass tube for each transformation, mixed with the ligation mixture and

incubated on ice for 30 minutes. The cells were then heat-shocked at 42°C for 2 minutes and returned to ice for 1 minute. 0.9ml of L-broth was added to each tube and the cells incubated at 37°C for 1 hour.

- b) Transformation in the presence of DMSO: The cells were pelleted by centrifugation (2 Krpm, 10 minutes at 4°C) and resuspended in 1/10th volume of Transformation Buffer which consists of L-broth containing 10% (v/v) PEG 3000, 5% (v/v) DMSO, 10mM MgCl₂ and 10mM MgSO₄. The cells were incubated on ice for 10 minutes. A 0.1ml aliquot of this suspension per transformation was added to each ligation mixture and incubated on ice for 30 minutes. 0.9ml of Transformation Buffer and 20μl of a 20% (w/v) glucose solution were added to each transformation, and the cells were allowed to grow at 37°C for 1 hour. Samples of 100-200μl from each transformation were then plated out on to L-agar plates supplemented with ampicillin (to a final concentration of 50μg/ml). The plates were inverted and incubated at 37°C overnight.

2.5.1.9 Preparation of Dialysis Tubing

Dialysis tubing (32mm flat width) was purchased from SIGMA. The tubing was cut into pieces of convenient length (10-20 cm), and boiled for 10 minutes in 2 litres of 2% (w/v) sodium bicarbonate/1mM EDTA. The tubing was then thoroughly rinsed in dH₂O followed by boiling in dH₂O for 10 minutes, allowed to cool and stored submerged in 70% (w/v) ethanol at 4°C. Prior to use the tubing was rinsed in dH₂O.

2.5.2 Preparation of Genomic DNA from Adults of *Drosophila melanogaster*

A rapid method for DNA extraction from single flies was used, essentially this method is that of Marcus (1985). Care was taken to prevent shearing of genomic DNA.

One or two flies were etherised, frozen in liquid nitrogen and homogenised in 50 μ l Genomic DNA Extraction Buffer (0.15M NaCl, 50mM Tris HCl pH 8.0, 15mM EDTA). The flies were homogenised in a 1.5 ml Eppendorf tube with a pasteur pipette, the tip of which had been melted into a ball to fit snugly in the bottom of the tube. A further 50 μ l of the same buffer containing 0.04% (w/v) SDS was added. Proteinase K was added to a final concentration of 50 μ g/ml and the homogenate incubated at 65°C for 30-60 minutes. The mixture was then phenol extracted twice and ethanol precipitated for at least 3-4 hours. The nucleic acid pellet was left to resuspend in TE buffer at 4°C overnight.

Genomic DNA was digested for 2 hours with 10 units of chosen restriction enzyme, in the appropriate buffer conditions (2.1), and RNaseA (to a final concentration of 0.1 μ g/ μ l) was also added to the reaction mixture.

2.5.3 Preparation of RNA from Different Developmental Stages and Tissues of *Drosophila melanogaster*

2.5.3.1 Large Scale Total RNA Preparations from Eggs, Larvae, Pupae and Adults

Usually up to 3g of eggs (0-6 hours), 4g of larvae (1st, 2nd, early third and late third instar), 5g of pupae, 2g of female adults and 2g of male adults were collected as described in section 2.3. Each sample was frozen in liquid nitrogen and homogenised in 5mls of RNA Extraction Buffer (100mM Tris HCl pH 7.5, 10mM EDTA, 150mM LiCl, 1% (w/v) SDS), in a 10ml sterile glass homogeniser (which had been previously baked at 200°C for 2 hours rendering it RNase free). The homogenate was extracted 4 times with a 1:1 ratio of phenol:chloroform (2.5.1.1) and then ethanol precipitated overnight at -20°C (2.5.1.3). The precipitate was pelleted by centrifugation in a Sorval RC-5B centrifuge (7 Krpm, 10 minutes), washed in 70% (w/v) ethanol and dissolved in 1-2 mls of sterile distilled water, and 2

volumes of 3M Lithium Chloride added. The sample was incubated at 4°C overnight, centrifuged and the pellet washed in 70% (w/v) ethanol, vacuum dried and resuspended in sterile distilled water.

2.5.3.2 Small Scale Total RNA Preparations from Ovaries, Testes and Carcasses

Up to 200 adults were hand dissected using watchmaker's forceps. The pairs of ovaries and testes and the remainder of the fly (carcass) were transferred to separate Eppendorf tubes which contained 100 μ l of RNA extraction buffer (2.5.3.1) and kept on ice. The samples were frozen in liquid nitrogen and homogenised in a 5ml sterile, RNase free glass homogeniser, in a total volume of 1-2mls RNA extraction buffer. Total RNA was extracted as above (2.5.3.1).

2.5.3.3 Preparation of poly (A)⁺ RNA

- a) Selection of poly (A)⁺ RNA using HybondTM-messenger Affinity Paper (mAP).

A piece of HybondTM-messenger affinity paper (mAP) was cut to 1cm² in size. This paper binds at least 25 μ g of poly (A)⁺ RNA/cm². Selection of poly (A)⁺ RNA from populations of total RNA was carried out according to the manufacturer's instructions (Amersham). The total RNA sample was heated to 65°C, cooled on ice and applied to the messenger affinity paper in high salt conditions (0.5M NaCl), which allows poly (A)⁺ RNA to bind to the mAP paper. Any unbound RNA is removed by washing with 0.5M NaCl followed by 70% (w/v) ethanol washes. The poly (A)⁺ RNA was then recovered by boiling the mAP paper in sterile dH₂O for 5 mins followed by ethanol precipitation (2.5.1.3).

- b) Selection of poly (A)⁺ RNA using oligo (dT) Cellulose Powder - Spin Method.

Oligo(dT) cellulose powder was purchased from Boehringer Mannheim, and the manufacturer estimates a binding capacity of 40mg of poly (A)⁺ RNA/g oligo(dT) cellulose powder. Total RNA was made up to a volume of 500μl with sterile dH₂O, heated to 65°C and cooled on ice. An equal volume of 2X Loading Buffer (40mM Tris HCl pH 7.6, 1M NaCl, 2mM EDTA, 0.2% (w/v) SDS) was added, and the RNA sample applied to 0.01g of dry oligo(dT) cellulose powder in an Eppendorf tube. The RNA sample and oligo(dT) cellulose powder were mixed by shaking vigorously for 5 minutes. The oligo(dT) cellulose powder was pelleted by centrifugation in a microfuge, for 5 minutes. The supernatant was removed, heated at 65°C for 5 minutes, cooled on ice and reapplied to the oligo(dT) cellulose powder as before. This procedure was repeated once more. Any unbound RNA was removed by washing 5 times with 1X Loading Buffer (20mM Tris HCl pH 7.6, 0.5M NaCl, 1mM EDTA, 0.1% (w/v) SDS). The poly (A)⁺ RNA was recovered from the oligo(dT) cellulose powder by adding 500μl of Elution Buffer (10mM Tris HCl pH 7.5, 1mM EDTA, 0.05% (w/v) SDS), followed by ethanol precipitation (2.5.1.3).

c) Estimation of poly (A)⁺ RNA Yields.

The poly (A)⁺ RNA yield was estimated to be 2% of the total RNA yield. Using both methods (see above) for poly (A)⁺ RNA extraction usually about 70% of the predicted poly (A)⁺ RNA yield was obtained. The OD₂₆₀ of a sample from a poly (A)⁺ RNA preparation was occasionally used as a guide for estimating the yield. The most reliable method of assessing the quantity and quality of a given poly (A)⁺ RNA preparation involved radioactively probing (2.5.8) Northern Blots (2.5.9) containing poly (A)⁺ RNA and measuring the strength and quality of the signal obtained compared to the

signal obtained from a track loaded with a known quantity of poly (A)⁺ RNA.

2.5.4 Preparation of Plasmid DNA

2.5.4.1 Small Scale

Small scale plasmid DNA was prepared by modification of the rapid extraction method of Birnboim and Doly (1979).

A single bacterial colony was used to inoculate 5ml of L-broth and grown with shaking, at 37°C overnight. A 1.5 ml aliquot of this culture was pelleted in a microfuge (12 Krpm, 1 min) and the supernatant discarded. The cells were resuspended in 100μl of 50mM glucose, 10mM EDTA, 15mM Tris.HCl pH 8.0 and 10μl of lysosyme (10mg/ml) were added, vortexed, and the mixture was incubated on ice for a further 2 minutes before the addition of 150μl of 3M sodium acetate pH 5.5 and vortexing. The sample was then phenol extracted (2.5.1.1), the DNA precipitated with ethanol (2.4.1.3) and the DNA pellet resuspended in 20μl TE buffer. Aliquots of 2μl were used in restriction digests (2.4.1.5) with RNase-A added to a final concentration of 0.1μg/μl in the reaction mixture.

2.5.4.2 Large Scale

A fresh overnight culture (as small scale) of bacteria containing the plasmid required was used to inoculate 500ml L-broth supplemented with the appropriate antibiotic and grown overnight, with shaking, at 37°C. The cells were pelleted in a Sorval RC-5B centrifuge (5 Krpm, 5 minutes) and resuspended in 10ml of ice cold 50mM glucose, 10mM EDTA, 25mM Tris HCl pH 8.0 with the addition of 200μl lysosyme (5mg/ml). After the suspension was incubated on ice for 5 minutes, 20ml of 0.5M NaOH, 1% (w/v) SDS were added, with vigorous mixing, and incubated on ice for 5 minutes after which 15ml of 3M Sodium Acetate pH 5.5 were added and the mixture incubated on ice for 60 minutes. After centrifugation (12 Krpm, 10 minutes), the supernatant was phenol extracted (2.5.1.1) and the DNA precipitated by

the addition of 0.6 volumes of propan-2-ol and incubation at -20°C for 60 minutes. The DNA was pelleted (12 Krpm, 20 minutes), resuspended in 10ml TE buffer and 10g Caesium Chloride and 1ml of Ethidium Bromide (10mg/ml) were added.

The resulting solution was centrifuged in a Sorval OTD 50B ultracentrifuge at 38 Krpm at 18°C, for 48 hours. The DNA bands were visualised using UV illumination and the plasmid band collected. Ethidium Bromide was removed by extraction with Caesium Chloride saturated butan-2-ol, the DNA precipitated with ethanol (2.5.1.3) and resuspended in TE buffer.

2.5.5 Manipulations with Lambda

2.5.5.1 Plating of Phages

Plating cells were prepared by growing host cells overnight in L-broth supplemented with 0.04% (w/v) maltose, pelleting the cells by centrifugation and resuspending in $\frac{1}{2}$ volume 10mM MgSO₄. Appropriate dilutions of phage were mixed with 0.1ml of plating cells and incubated at 37°C for 15 minutes. 3ml of molten BBL top agar was added and poured onto BBL agar plates (9cm diameter). The plates were allowed to set then inverted and incubated at 37°C for 6-16 hours. Titres of phage suspensions were calculated from the number of plaques produced by a given dilution.

2.5.5.2 Plate Lysates - Phage Storage

A well isolated phage plaque was picked into 1ml of phage buffer and left to elute from the agar plug at 4°C overnight. The next day, 50μl of phage suspension was incubated with 100μl of plating cells (prepared as in section 2.5.5.1) and poured onto L-agar plates (9cm diameter). Phage were grown non-inverted for 12-16 hours at 37°C, until the surface of the plate was confluent. The phage were harvested by adding 5ml of phage buffer onto the surface of the plate and incubating for 2-3 hours at 4°C. The liquid layer was removed, vortexed briefly and cells

and agar removed under centrifugation (Sorvall RC5-B, 5 Krpm, 5 minutes). The supernatant was removed, titred and stored at 4°C with the addition of a few drops of chloroform.

2.5.5.3 Phage DNA Extraction

a) Miniscale

A single plaque was picked and transferred into 0.3ml of adsorption buffer (10mM MgCl₂, 10mM MgSO₄), and left at 4°C overnight. The next day, 1ml of host plating cells (see section 2.5.5.1) was added to the phage suspension and incubated at 37°C for 15 minutes. The cells and phage were added to 10mls of L-broth in a flask which was supplemented with 10mM MgCl₂ and 0.1% (w/v) glucose, and the phage allowed to lyse the host cells by shaking at 37°C for 3-8 hours. After lysis had taken place, the cell debris was removed by centrifugation (10 Krpm, 30 minutes). The supernatant was carefully removed to a fresh sterile flask and DNase I and RNase A were added, each to a final concentration of 1µg/ml, and incubated at 65°C for 30-60 minutes. The phage were pelleted by ultracentrifugation (Sorvall OTD-50B, 30 Krpm, 60 minutes) and resuspended in 200µl of TM buffer (50mM Tris HCl pH 7.5, 10mM MgSO₄). Proteinase K was added to the resuspended phage pellet to a final concentration of 1µg/µl and incubated at 65°C for 1 hour. Phage DNA was extracted by gentle shaking with an equal volume of phenol (section 2.5.1.1), followed by a phenol : chloroform (1 : 1) extraction and finally traces of phenol were removed by extraction with an equal volume of chloroform. The λ DNA was then ethanol precipitated (2.5.1.3), and resuspended in TE buffer for several hours at 4°C and stored at -20°C.

b) Maxi Scale

Host cells were grown overnight in a 5ml culture of L-broth supplemented with 0.04% (w/v) maltose. A 0.5ml sample was removed and the cells pelleted by centrifugation in a microfuge and resuspended in 0.5mls of phage buffer. To this preparation of cells, 0.5ml 10mM MgCl₂ and a 0.5ml suspension of phage containing 2.5×10^8 plaque forming units (pfu) were added, and incubated at 37°C for 15 minutes. The cells and phage were then added to 200mls of L-broth supplemented with 10mM MgCl₂ in a 2 litre flask and grown for 6-7 hours, shaking at 37°C until lysis was apparent by the presence of cell debris. Chloroform (0.5mls) was then added to complete lysis. The cell debris was pelleted at 10 Krpm for 10 minutes, and the supernatant was removed and treated with RNaseA and DNase (both at a final concentration of 1μg/ml) at room temperature for 1-2 hours. The degraded cellular RNA and DNA was removed by centrifugation (10 Krpm, 10 minutes). The phage was precipitated by adding 8g of NaCl and 8g of polyethylene glycol 6000 and incubating at 4°C overnight. The phage were pelleted by centrifugation (7 Krpm, 15 minutes), and resuspended in 8mls of phage buffer and extracted once by gentle shaking with an equal volume of chloroform. The phage particles were further purified on a CsCl step gradient. Three 2ml CsCl steps (in phage buffer) were used with densities of 1.3, 1.5 and 1.7g/ml. The phage sample was carefully layered onto the step gradient in 14ml MSE centrifuge tubes and the phage banded at 35 Krpm at 20°C for 2 hours in a 6 x 14ml Ti swing out rotor (Sorvall Ultracentrifuge OTD 50B). The bluish phage band was collected through the side of the tube with a syringe. The phage suspension was then dialysed twice for 1 hour at room temperature against 1 litre of 10mM MgCl₂, 10mM NaCl, 50mM Tris HCl pH 8.0 to remove CsCl. Dialysis tubing was prepared in the standard way (2.5.1.10). The dialysed phage

suspension was diluted to 2ml with TE and 100 μ l 0.1M EDTA, 100 μ l 10% (w/v) SDS and 10 μ l 20mg/ml proteinase K added and incubated at 65°C for 1 hour. The phage DNA was then deproteinised by phenol extraction (2.5.1.1) and ethanol precipitated (2.5.1.3). The DNA pellet was left to dissolve in 50 μ l of TE buffer overnight at 4°C before storing at -20°C.

2.5.5.4 Screening libraries / Plaque Lifts

DNA from phage plaques were transferred to Hybond-N nylon membrane filters (Amersham). For the primary screens approximately 1×10^5 plaques were plated on large square petri dishes using the appropriate plating cell strain. Phage were combined with plating cells and placed at 37°C for 15 minutes. The infected cells were added to 25 mls of 0.7% agarose, 10mM Mg₂SO₄ in L-broth at 49°C and the mixture was immediately poured on to the surface of a dry 23 X 23 BBL bottom (λ EMBL 4, λ DASH II, λ gt10), L-amp bottom (λ gt11), or Tetracycline (10 ug/ml- κ ZAP II) 23 X 23 cm plate and allowed to grow for 6-12 hours at 37°C. After cooling at 4°C for 1 hour a single sheet of Hybond-N nylon membrane (circles 9cm in diameter or 20cm x 20cm sheets for library lifts-Amersham) was laid on the surface for 1 minute and marked with needle holes for subsequent orientation. The membrane was removed and placed plaque side up on blotting paper soaked in 0.5M NaOH, 1.5M NaCl for 7 minutes then neutralised by placing on blotting paper soaked in 1M Tris HCl pH 8.0, 1.5M NaCl for 6 minutes, and finally placed on blotting paper soaked in 2 X SSC for 5 minutes. The filters were air dried for 30 minutes and the DNA fixed to the membrane using ultra violet light as described in section 2.5.7.1. Secondary screens were performed to permit purification of positive phage clones, using small plates (90mm) and a phage density not greater than 250 pfu/plate. When an isolated positive clone had been identified it was picked with a capillary tube and

placed in 1ml of phage buffer containing a drop of chloroform. The phage stock was stored at 4°C.

2.5.5.5 In Vivo excision of pBluescript plasmid from λ ZAP II

The plaque of interest was cored from an agar plate and transferred to an eppendorf tube containing 500 ul of phage buffer and 20 ul of chloroform. The tube was left for 2 hours at R.T. to allow phage to elute.

In a 50 ml conical tube were combined 200ul of O.D.₆₀₀=1.0 XL-1 Blue cells, 200 ul of λ ZAP II phage stock ($\geq 10^5$ phage particles) and 1 ul of R408 helper phage ($\geq 10^6$ pfu / ml). The mixture was incubated at 37°C for 15 minutes then added to 5 ml of 2 X YT medium and incubated for 3 hours with shaking. The tube was heated for 20 minutes at 70°C then spun for 5 minutes at 4000g to pellet cells. The supernatant contained pBluescript phagemid packaged as a phage particle.

To plate the excised phagemid 10 ul of phage stock was combined with 200 ul of OD₆₀₀=1 XL-1 Blue cells the infected cells were plated on L-Amp plates and incubated overnight to yield colonies.

2.5.6 Agarose Gel Electrophoresis

2.5.6.1 Agarose Gel Electrophoresis of DNA

Agarose gels were made and run in 1 X TBE gel buffer. Ethidium Bromide was added at a concentration of 0.5mg/ml to both the gel and the running buffer. Gel loading sample buffer was added to the DNA samples to 10% (v/v).

Mini-gels were cast and run in a Cambridge Biotechnology model CB1000 mini-gel apparatus. These gels were run at 50-100V for 1-2 hours. Larger agarose gels (16 x 10 x 1cm) were run submerged in home-made gel kits at 100-150V for 1-4 hours.

After electrophoresis DNA bands were visualised on a chromatovue C-70G UV viewing system (254nm short, 265 long).

2.5.6.2 Agarose Gel Electrophoresis of RNA

RNA samples (typically 10 μ g total RNA in dH₂O or 2 μ g of poly (A)⁺ RNA in dH₂O) were added to an equal volume of formaldehyde sample buffer (FSB) and heated to 60°C for 5 minutes, then immediately cooled on ice before adding 0.25 volumes of Ficoll-dye-EDTA (FDE). The samples were immediately loaded on to 1.3% (w/v) agarose gels which were made and run in 1X MOPS buffer. The gels also had formaldehyde added to 17.3% (v/v). The gels (16 x 10 x 1cm) were run submerged in home-made gel kits at 100 volts for 3-4 hours.

2.5.6.3 Molecular Weight Markers

DNA restriction fragments were sized by comparing their distance migrated in the gel with the migration distances of DNA fragments of known molecular weight. λ DNA digested with HindIII DNA (purchased from NBL) and a set of DNA fragments differing in size by 1kb (NBL) were used as standard molecular weight markers.

2.5.7 Transfers to Membrane Filters

2.5.7.1 Southern Blotting (Southern, 1975)

Gels for Southern transfers were prepared with Pharmacia NA agarose. *Drosophila* genomic DNA samples were run in 1% (v/v) agarose gels. Plasmid and λ DNA restriction fragments were sized on 0.7 -1.2% (v/v) agarose gels. When electrophoresis was complete, and after photographing, gels were gently shaken for 30 minutes in 0.5M NaOH, 1.5 NaCl then soaked for 30 minutes in 1M Tris HCl pH 5.5, 3M NaCl and washed in 2X SSC.

A blotting paper wick was placed on a platform supported above a reservoir of 20X SSC with the ends of the

wick immersed in the buffer. The wick was thoroughly wet with 20X SSC and the gel laid on top. A sheet of Hybond-N membrane filter (Amersham) was cut to the same dimensions as the gel and was wetted with 2X SSC before being laid on top of the membrane. Care was taken during these steps to avoid air bubbles being trapped. A few pieces of blotting paper, cut to the same dimensions as the gel were soaked in 2X SSC and placed on top of the membrane. Further pieces of blotting paper (again cut to the same dimensions as the gel) were laid, dry, on top of the wet blotting paper to a thickness of 3-5cm, followed by a stack of paper towels to a thickness of 5cm. A sheet of perspex was laid on top of the paper towels and a weight placed on top. Transfer was allowed to occur for 16-24 hours after which the membrane filter was removed, rinsed briefly in 2X SSC and wrapped in clingwrap (Vitafilm, Goodyear Tyre and Rubber Co. Ltd, Staffordshire, GB). The DNA was fixed to the filter by UV irradiation in a UV light box (ChromatoVue C-70G, UV viewing system), with the DNA side towards the source of UV light (254nm short and 365nm long), for 5 minutes. The membrane was sealed and stored in a plastic bag.

2.5.7.2 Bidirectional Southern Blotting

Some DNA gels for Southern transfers were blotted onto two sheets of Hybond-N nylon membrane filters. After gel electrophoresis, the gel was denatured for 30 minutes in 0.5M NaOH, 1.5 M NaCl then soaked for 30 minutes in 1M Ammonium Acetate/0.02M NaOH for 1 hour. Two sheets of Hybond-N membrane were cut to the same size as the gel and soaked in 1M Ammonium Acetate/0.02M NaOH for 5 minutes. To set up the blot a stack of paper towels 3-5cm in height were cut to the same size as the gel and 3 sheets of blotting paper moistened with 1M Ammonium Acetate/0.02M NaOH were placed on top. Next, a sheet of moistened Hybond-N membrane (with 1M Ammonium Acetate/0.02M NaOH) was laid on top of the blotting paper and the gel placed on top. Another piece of Hybond-N membrane was placed on the

upper side of the gel, so that the gel was now sandwiched between 2 sheets of transfer membranes. The blot was stacked up by adding moistened (as before) blotting paper followed by 3-5cm of paper towels. A weight was placed on top of the blot and DNA transferred equally onto each membrane after a few hours. The two membranes were removed and the DNA fixed using UV light (2.5.7.1).

2.5.7.3 Northern Blotting (Thomas, 1980)

Gels for Northern transfers were prepared (2.5.6.2) with Miles Seachem agarose. After electrophoresis, and without further treatment, the gels were transferred to a platform with a blotting paper wick in a reservoir of 20X SSC. The remainder of the transfer procedure was carried out as described for Southern blotting (2.5.7.1).

2.5.8 Radioactive Labelling of DNA and Autoradiography

2.5.8.1 Radioactive Labelling of DNA by Random Priming (Feinberg and Vogelstein, 1983, 1984)

The DNA fragment to be labelled was denatured by boiling for 5 minutes followed immediately by rapid cooling on ice. If the DNA fragment was isolated in low melting point agarose, the fragment was denatured by boiling for 5 minutes, allowed to cool for 1 minute at room temperature then added directly to the reaction mixture. Between 30 - 100ng of DNA was labelled in the reaction in a total volume of 50 μ l containing 10 μ l OLB (detailed in section 2.1), 1 μ l 20mg/ml BSA, 20-50 μ Ci [³²P]-dCTP, 1 μ l (5 units/ μ l) Klenow fragment of DNA polymerase I. The reaction was incubated for 4-16 hours at room temperature.

2.5.8.2 Measurement of Radioactivity Incorporated into DNA

The efficiency of incorporated radiolabel was measured by precipitation with tri-chloroacetic acid (TCA). One aliquot of 1 μ l from the reaction mixture was added to 200 μ l

ice cold dH_2O containing $10\mu\text{g}$ BSA and incubated on ice for 10 minutes. The sample was made 10% (w/v) TCA and the precipitate collected on 2.5 cm Whatman GF/C glass fibre circles using a Buchner funnel equipped with a water pump. The discs were washed with 5% (w/v) TCA and dried. Another sample ($1\mu\text{l}$) was spotted directly on to a Whatman GF/C glass filter and dried. Each filter was inserted into a scintillation vial, immersed in scintillant (6g/L butyl-PBD in toluene) and the radioactivity measured by counting in a liquid scintillation spectrometer (Intertechnique SL 3000). The percentage incorporation of label into nucleic acid was estimated by comparing the two values and usually probes used achieved 50-90% incorporation and were 10^7 - $10^9\text{cpm}/\mu\text{g}$ DNA.

2.5.8.3 Precipitation of Radiolabelled DNA Molecules

Radiolabelled DNA fragments were precipitated from the labelling reaction by the addition of an equal volume of 4M Ammonium Acetate (pH 6.6) and 4 initial volumes of ethanol. After incubation in a dry ice/methanol bath for 15 minutes, the DNA fragments were pelleted by centrifugation in a microfuge for 5 minutes, washed in 70% (w/v) ethanol, air dried and resuspended in hybridisation buffer (see section 2.8). This precipitation is successful in removing >90% of unincorporated labelled triphosphates from labelling reactions (Okayama and Berg, 1982).

2.5.8.4 Autoradiography

Autoradiography of nylon membrane filters was performed using Cronex 4 (Dupont) X-ray film and cassettes @ -70°C . Films were developed in an Agfa 1 automatic film processor. The procedure for autoradiography of sequencing gels was identical except exposure was at room temperature.

2.5.8.5 RNA probes

DIG-11-UTP or ^{32}P UTP labelled RNA probes were prepared from recombinant plasmid (p-GEM-1) or λ (λ Dash II) templates respectively. Plasmid templates were first linearised with an appropriate restriction enzyme recognising a unique site within the polylinker. λ clones were cut with a restriction enzyme that cut within the cloned insert to allow the generation of an "end-specific" probe. RNA polymerase (SP6 or T7 was obtained from Boehringer Mannheim). All solutions were made up with DEPC treated water.

Template DNA (1 μg of plasmid or 5 μg of λ) was added to an eppendorf tube containing 2 μl of 10 x NTP mix (ATP, GTP and CTP @ 10 mM each; UTP @ 6.5 mM and DIG-11-UTP @ 3.5 mM in Tris-HCl neutralised solution pH 7.5), 2 μl of 10 x transcription buffer (0.4M Tris-HCl pH 8.0; 60 mM MgCl_2 ; 100 mM DTT; 20 mM spermidine; 100 mM NaCl; 1 U/ μl RNAase inhibitor), 40 U of RNA polymerase. The reaction volume was made up to 20 μl with DEPC treated H_2O and the tube was incubated at 37°C for two hours. Unincorporated nucleotides were removed as described in section 2.5.8.3. DIG-11-UTP labelled plasmid derived probes were used for *in situ* hybridisation to polytene chromosome squashes (section 2.11) and λ DASH II derived probes were used for genomic library screening (section 2.5.5.4).

2.5.9 DNA sequencing by the Dideoxynucleotide chain termination method (Sanger et.al., 1977)

Template preparation

Sequence data was obtained from both single and double stranded templates. Single stranded templates were prepared as follows:

2.5.9.1 M13

Recombinant phage plaques were picked from minimal medium plates into 1.5 ml of TY broth containing 15 μ l of a fresh overnight of TG1 cells. The culture was allowed to grow for 5.5 hours, transferred to an eppendorf tube and spun for 10 minutes at R.T. to remove the cells. The supernatant was transferred to a new tube, spun for a further 10 minutes then transferred to a third tube. 150 μ l of 20% (w/v) PEG 8000, 2.5 M NaCl were added and the tube was incubated for 15 minutes at R.T. then spun for 15 minutes. The supernatant was carefully removed and discarded and the pellet of precipitated phage particles was resuspended in 100 μ l of TE. The phage suspension was phenol extracted twice with 50 μ l of phenol (section 2.5.1.1) and the aqueous layer was ethanol precipitated (section 2.5.1.3). The DNA pellet was dissolved in 30 μ l of TE. 10 μ l of this solution was used per sequencing reaction.

2.5.9.2 pBluescript - single strand rescue

A single ampicillin/tetracycline resistant colony was innoculated into 5ml of 2X YT broth containing 75 μ g/ml ampicillin and VCS (kanamycin resistant helper phage) at 10^7 to 10^8 pfu/ml (multiplicity of infection approximately 10). The culture was grown at 37°C with vigorous aeration for 1-2 hours and kanamycin was added to 70 μ g/ml. After a further 16-24 hours of growth 1.5 ml aliquots of the culture were transferred to eppendorf tubes. Subsequent steps in the procedure were as for M13 (section 2.5.9.1).

2.5.9.3 Double stranded templates

Double stranded templates were prepared as follows: Before annealing the sequencing primer to a double stranded template it is necessary to denature it. The plasmid or pBluescript phagemid DNA was obtained from a maxi-prep as

described in section 2.5.4.2. The amount of DNA required for each sequencing reaction was dissolved in 8 μ l of dsH_2O and added to an eppendorf tube containing 2 μ l of 2 M NaOH. After a 15 minute incubation at R.T. 3 μ l of 3 M sodium acetate (pH 4.8), 7 μ l of dsH_2O and 60 μ l of chilled absolute ethanol were added to the denatured template and the mixture was incubated on dry ice for 15 minutes. The DNA was pelleted by spinning for 15 minutes and the supernatant removed. The DNA pellet was washed in 70% ethanol and resuspended in 10 μ l of dsH_2O for use in a sequencing reaction.

2.5.9.4 Sequencing reactions

a) Annealing

The primer was annealed to the single stranded template by adding 2 μ l of annealing buffer (0.1 M Tris-HCl pH 8.0, 50 mM CaCl_2) and 2 μ l of primer at a concentration of 80 μ M. The mixture was spun down and incubated at 65°C for 10 minutes after which it was placed at R.T. for a minimum of 10 more minutes.

The denatured double stranded DNA was identically treated except that the DNA/primer mixture was incubated at 37°C for 20 minutes before the 10 minute incubation at R.T.

b) Extension/Termination reactions

Annealed templates (single or double stranded) were treated in the same manner. The tubes containing the annealed templates were placed in a rack and 1 μ l of 0.1 M DTT, 1 μ l of labelling mix (10 mM DTT, 0.25 mM of each dNTP), 2 μ l of T₇ DNA polymerase at 1.5 U/ μ l and 1 μ l of ^{35}S -dATP (10 μ Ci/ μ l) were added. The tube was spun and incubated at R.T. for 5 minutes. To each of four tubes previously labelled T,C,G,A and containing 2.5 μ l of T,C,G and A termination mix respectively (Table 2.5), 2.5 μ l of

Table 2.5 Composition of ddNTP solutions (all volumes in microliters)

	T mix	C mix	G mix	A mix
0.5 mM dTTP	25	500	500	500
0.5 mM dCTP	500	25	500	500
0.5 mM dGTP	500	500	25	500
10 mM ddTTP	40	-	-	-
10 mM ddCTP	-	6	-	-
10 mM ddGTP	-	-	16	-
10 mM ddATP	-	-	-	0.6
TE	935	969	959	500

incubated at 37°C for 5 minutes. The reactions were stopped by adding 5 μ l of 'stop mix' (0.3% (w/v) Xylene cyanol, 0.3% (w/v) Bromophenol blue, 10 mM EDTA in 100% deionised formamide). The DNA strands were denatured by boiling for 2 minutes, immediately before loading on to a polyacrylamide sequencing gel (section 2.5.9.5).

2.5.9.5 Polyacrylamide gel electrophoresis for sequencing

Sequencing gels (40 cm x 30 cm x 0.4 mm; 45 cm x 18 cm x 0.2 mm or 60 cm x 18 cm x 0.2 mm) were 6% (w/v) acrylamide (19:1 acrylamide : bis-acrylamide) in TBE buffer containing 7.7 M urea. Sequencing gels were run at a constant power of 40 watts for plates 40 cm x 30 cm x 0.4 mm, for times ranging between 2 and 9 hours. For plates 45 cm x 18 cm x 0.2 mm the power used was 35 watts and for plates 60 cm x 18 cm x 0.2 mm the power used was 45 watts, both run for 2 to 9 hours. After electrophoresis, gels were fixed in 10% (v/v) acetic, 10% (v/v) methanol for 15-20 minutes, transferred onto blotting paper (Ford Goldmedal Blotting Paper), dried on a gel drier (Biorad model 583) at 80°C for 2 hours and autoradiographed (section 2.5.8.4) for 14-20 hours at room temperature.

2.6 Construction of genomic libraries from transposants K8 and A22

Genomic DNA was extracted from transposants K8 and A22 by the method described in section 2.5.2. DNA of the λ replacement vector EMBL 4 was prepared according to the method described in section 2.5.5.3 b. 5 μ l of the *Drosophila* DNA and 5 μ l of vector DNA were combined, ethanol precipitated (section 2.5.1.3) and resuspended in 7.5 μ l of ds H₂O. For the ligation reaction, 1 μ l of 10 X ligation buffer, 1 μ l of 10mM ATP and 0.1 U of T4 DNA ligase were added to the DNA solution. The tube was spun, incubated at 4°C overnight and the ligation reaction stopped by heating at 70°C for 5 minutes. The tube was then stored on ice. The *in vitro* packaging extracts (obtained from Amersham International) were slowly thawed on ice. Immediately upon thawing of the extracts, the DNA was added to extract A (prepared from *E.coli* strain BHB2688) then extract B (prepared from *E.coli* strain BHB2690) was added and the tube was gently mixed. The tube was spun and incubated at

R.T. for 2 hours. After the incubation, 0.5 ml of phage buffer and 10 μ l of chloroform were added. The phage stock was stored at 4°C and subsequently titred (section 2.5.5.1). The library was screened using *E.coli* strain NM646 which carries a P2 lysogen to select against non-recombinant phage.

2.7 Polymerase chain reaction

The polymerase chain reaction (PCR) was performed using Cetus Taq polymerase and Cetus Taq polymerase reaction buffer (10 x buffer contains 500 mM KCl, 100 mM Tris-HCl pH 8.3, 15mM MgCl₂ and 0.01% (w/v) gelatin).

The reaction mix was as follows:

dNTP mix (dATP,dTTP,dGTP,dCTP each @ 40mM)	2 μ l
10 x Cetus Taq polymerase reaction buffer	10 μ l
Primer oligonucleotide no.1	10 μ l
Primer oligonucleotide no.2	10 μ l
genomic DNA (30 ng/ μ l)	3 μ l
dH ₂ O	63 μ l
<hr/>	
Total	98 μ l

The above reagents were combined in an eppendorf tube and overlayed with 100 μ l of filter sterile mineral oil saturated with 1 X Cetus Taq polymerase buffer to reduce evaporation. The tube lid was pierced and in order to maintain a good thermal contact, the bottom of the tube was smeared with a high vacuum silicone grease.

The reaction was performed in a Hybaid programmable heating block, according to the method of Kaiser and Goodwin (1990). After a period of 5 minutes at 95°C to denature the template, 2 μ l of Taq polymerase @ 1U/ μ l was added to the reaction mixture. The reaction was then taken through 30 cycles (1 minute at 50°C; 3 minutes at 70°C and

1 minute at 95°C) and finally incubated for 5 minutes at 50°C and 20 minutes at 70°C.

2.8 Hybridisation of Northern Blots, Southern Blots and Plaque Lifts

Hybond-N nylon membranes used for Northern blots, Southern blots and Plaque lifts were prehybridised in hybridisation solution (0.5M Sodium Phosphate pH 7.0, 7% (v/v) SDS, 50% (v/v) Formamide) in sealed plastic bags at 42°C for a minimum of 2 hours but usually filters were pre-hybridised overnight. A volume of 10mls of hybridisation solution was used for most blots (16cm x 10cm) and plaque lifts (9cm diameter). For library lifts (20cm x 20cm) 20mls of hybridisation solution was used.

Once the radiolabelled probe had been denatured (90°C, 5 minutes) it was added to the bag containing the membrane and hybridisation solution and the probe allowed to anneal at 42°C overnight.

2.8.1 Post-Hybridisation Washing

After the probe had been removed, membrane filters were washed at varying levels of stringency as follows.

a) Northern Blots

Northern blots were washed in two changes of 1X SSC 0.5% (v/v) SDS at 42°C for 1 hour each and exposed to X-ray film (2.5.8.4). If an overnight exposure detected non-specific background on the filter, a further wash in 1X SSC, 0.5% (v/v) SDS at 65°C for 1 hour was usually sufficient to remove most background.

b) Southern Blots and Plaque Lifts

Genomic blots were washed in two changes of 1X SSC, 0.5% (v/v) SDS at 42°C for 1 hour each followed by a further wash in the same solution at 65°C for 1 hour and exposed to X-ray film (2.5.8.4). Plasmid DNA and λ DNA Southern filters and Plaque lifts were washed briefly (15 minutes) at 42°C in 1X SSC, 0.5% (v/v) SDS followed by 2 washes in the same solution at 65°C for 1 hour each. Exposure to X-

ray film allowed signal detection usually within 1-2 hours if probed with homologous DNA.

2.9 Radioactive Labelling of poly (A)⁺ RNA by Reverse Transcription

Transcription of poly (A)⁺ RNA into cDNA was performed according to the method of Gubler and Hoffmann (1983) using a cDNA synthesis kit purchased from Boehringer Mannheim. At least 1 μ g of poly (A)⁺ RNA was added to a 20 μ l labelling reaction containing first strand synthesis buffer, RNaseA inhibitor, dGTP, dTTP, primer oligo(dT)₁₅, redistilled sterile water and 50 μ Ci of [³²P]-dCTP, 50 μ Ci of [³²P]-dATP and AMV reverse transcriptase. The reaction was incubated at 42°C for 1 hour and the amount of synthesis of [³²P]-dCTP and [³²P]-dATP labelled cDNA was measured by precipitation with TCA (2.5.7.2). Usually ~3 μ g of labelled cDNA was synthesised from 1 μ g of poly (A)⁺ RNA. Removal of unincorporated nucleotides was carried out as described previously (2.5.8.3). The labelled cDNA was resuspended in 1ml of sterile dH₂O.

2.9.1 Probing Southern Blots with Radiolabelled cDNA

Radiolabelled cDNA made from poly (A)⁺ RNA prepared from male and female adults of *Drosophila melanogaster* (2.5.3) was used to probe a set of restriction endonuclease digested recombinant λ bacteriophage clones which had been bidirectionally Southern blotted (2.5.7.2). The membranes were prehybridised in a sealed plastic bag containing hybridisation solution for 24 hours. The cDNA probe (2.5.11) was denatured at 90°C for 5 minutes, snap cooled on ice for 5 minutes and added to the hybridisation bag. The hybridisation was carried out at 42°C overnight. The probe was removed and the filter washed in 3 changes of 0.1X SSPE, 0.1% (v/v) SDS for 1 hour each wash. The excess moisture on the membrane surface was removed with blotting paper and the filter then exposed to pre-flashed Hyperfilm (Amersham).

2.10 In situ Hybridisation to Third Instar Polytene Chromosome Preparations with [3H]-dCTP labelled probes

2.10.1 Preparation of Third Instar Polytene Chromosomes

Microscope slides were washed in 1M HCl and then subbed by dipping into a solution of 0.1% (w/v) gelatin, 0.025% (w/v) chromic Potassium Sulphate, and air dried. Cover slips were washed in 1M HCl, air dried and siliconised by dipping into a solution of Dimethyldichlorosilane (about 2% (v/v) in 1,1,1-trichloroethane) followed by air drying.

Salivary glands were dissected in Ringer's Solution (2.1) from late third instar larvae and placed in a small drop of aceto-orcein stain (BDH) for five minutes. A few drops of 45% (w/v) acetic acid was added to allow the salivary gland chromosomes to be easily visualised. The glands were then transferred to a drop of 45% acetic acid on a microscope slide. A coverslip (18mm²) was placed on top of the glands. The cells were broken by pressing the coverslip at one edge with a tissue followed by gentle tapping with a stiff, pointed probe starting over the glands and then gently working towards the edge in a spiral pattern. The slide was then turned over, onto a paper towel, and pressed at one edge so that the stain flowed to the opposite edge thus separating cell debris from the nuclei. The slide was inverted again, held at the edge with a paper towel and the chromosomes were spread by streaking across the coverslip in a zig-zag motion using a pointed probe. Finally the chromosomes were flattened by placing the slide between filter paper circles and pressing firmly. The slide was frozen in liquid nitrogen, the coverslip removed, and then the chromosomes were dehydrated in 95% (w/v) ethanol for 5 minutes, and air dried. The chromosomes were then examined under phase contrast microscopy and chosen for subsequent hybridisation experiments if the chromosome morphology was in good

condition The slides were stored at 4°C in an air tight dry plastic container.

2.10.2 Prehybridisation of Polytene Chromosome

Preparations

Before hybridisation, the slides were incubated in 2X SSC at 65°C for 30 minutes, then dehydrated through 2 changes of 70% (w/v) ethanol, 10 minutes each and 1 change of 95% (w/v) ethanol for 5 minutes and air dried. 200µl of 100µg/ml RNaseA in 2X SSC was pipetted onto the chromosomes and a large coverslip (25mm²) was placed on top. The slides were incubated at 37°C in a moist chamber, then rinsed in 2X SSC (3 changes at 5 minutes each) followed by dehydration through an ethanol series as described above.

The chromosomal DNA was then denatured by soaking the slides in 7mM NaOH for 3 minutes followed by dehydration as before.

2.10.3 Preparation of [³H]-dCTP Labelled DNA Probes and Hybridisation

Probe DNA was labelled using the random priming method. The desired fragments were isolated from low melting point agarose gels (2.5.1.6). 50ng to 100 ng of DNA in low melting point agarose was labelled as described in section 2.5.8.1 using [³H]-dCTP instead of [³²P]-dCTP. The reaction components were added to 40µCi of [³H]-dCTP which has to be dried before use as it is supplied in 50% (v/v) ethanol. Labelled probe DNA was separated from unincorporated [³H]-dCTP by the addition of 50µl of 4M Ammonium Acetate pH 6.6 and 200µl of absolute ethanol and placed on a dry ice/methanol bath for 30 minutes. 10µg of sonicated salmon sperm DNA was also added to the ethanol precipitation. The labelled DNA was pelleted by centrifugation in a microfuge for 7 minutes, washed in 70% (w/v) ethanol, vacuum dried and resuspended in hybridisation buffer (50% (w/v) formamide, 4X SSC, 1X

Denhardt's). A volume of 10 μ l to 20 μ l of probe with approximately 5X 10⁵ cpm was pipetted onto each squash, and covered with an acid washed coverslip. The slides were sealed with cow gum (diluted with diethyl ether to a consistency easily pipetted using a glass pasteur pipette) and placed in a moist chamber at 42°C overnight.

2.10.4 Washing Steps After Hybridisation

After hybridisation the cow gum seal was easily removed using tweezers and the microscope slide floated off in the first wash. Slides were washed twice in 2X SSC, 50% (v/v) formamide for 10 minutes, at room temperature, then dehydrated through ethanol as described above (2.5.13.2).

2.10.5 Autoradiography

Prior to dipping, 5ml of 2% (w/v) glycerol was measured out into a dipping chamber and clamped immersed in a water bath at 42°C, in the darkroom. In complete darkness, approximately 5ml of Ilford IL4 photographic emulsion was melted in a beaker pre-clamped in the water bath, and then added to the glycerol in the dipping chamber. The slides were dipped one at a time, and the excess emulsion was removed by vigorous shaking. The slides were stored in a lightproof box for approximately 10 days at 4°C. The slides were developed in complete darkness for 3 minutes, stopped with 3% (w/v) acetic acid, fixed in Hypon solution for 1½ minutes, and washed in running water for 10 minutes, followed by air drying. The developer was 0.22% (w/v) p-methylamino phenol sulphate, 7.2% (w/v) NaSO₃, 0.88% (w/v) hydroquinone, 4.8% (w/v) Na₂CO₃, 0.4% (w/v) KBr. This can be purchased from Kodak (D19 developer).

2.10.6 Microscopy and Chromosome Staining

After developing, the slides were examined using light microscopy to determine if the hybridisation signal visualised as black grains at the site of hybridisation on

the chromosome arm were detectable. The slides were then stained as follows. The slides were soaked in 10mM Sodium Phosphate buffer, pH 7.0, for 3 minutes, then transferred to a 5% (v/v) dilution of Giemsa stain (supplied by SIGMA) in 10mM Sodium Phosphate buffer pH 7.0. Slides were stained for 10 minutes, then placed with 10mM Sodium Phosphate, pH 7.0 for 5 minutes and finally rinsed extensively in running water. The slides were air dried and viewed at X25, X40 and under immersion oil at X100 magnification to locate the precise band within which the signal was present. The *in situ* hybridisation signals were photographed using both black and white film (Ilford Pan F) and also colour film (Kodak Ektachrome 160).

2.11 in situ hybridisation to polytene chromosome preparations with non-radioactively labelled probes

The procedure for non-radioactive *in situ* hybridisation to chromosome squashes was identical to the above as far as the pre-hybridisation stage. The hybridisation solution contained Digoxigenin-11-uridine-5'-triphosphate (DIG-11-UTP from Boehringer mannheim) labelled RNA probe at a concentration of approximately 25 ng/ul. After the post-hybridisation washes the following immunodetection procedure was followed.

- 1) 2 X 3 minute washes in Buffer 1 (100 mM Tris-HCl, 150mM NaCl, pH 7.5)
- 2) 30 minute incubation in Buffer 2 (Buffer 1 containing 0.5 % (w/v) blocking reagent (Boehringer mannheim))
- 3) 5 minute wash in Buffer 1
- 4) 60 minute incubation at room temperature with gentle shaking in Buffer 1 containing 1:5000 fold dilution of alkaline phosphatase conjugated anti-digoxigenin antibody (Boehringer mannheim)
- 5) 2 X 15 minute washes in Buffer 1

6) 2 X 15 minute incubations in Buffer 3 (100 mM Tris-HCl, 100mM NaCl, 50 mM MgCl₂, pH 9.5)

7) Incubation for 15-30 minutes in colour detection solution [Buffer 3 containing .45 % NBT (4-Nitro blue tetrazolium chloride) and .35 % X-phosphate (5-Bromo-4chloro-3indolyl-β-D-galactopyranoside)(w/v)]

Colour development was monitored under the microscope at X 25 magnification and the chromogenic reaction was stopped by 4 X 15 minute washes in Buffer 4 (10 mM Tris-HCl pH 8). Subsequent microscopy and chromosome staining were as described above.

2.12 *In situ* Hybridisation To Whole Mounts of *Drosophila* Ovaries and Testes

(Tautz and Pfeifle, 1989)

2.12.1 Preparation of Tissues - Paraformaldehyde Fixation

Ovaries and testes were hand dissected from adults of *Drosophila*. The method of fixing is essentially the same method described in Tautz and Pfeifle (1989). The flies were anaesthetised and dissected in fixing solution which consists of 0.1M Hepes pH 6.9, 2mM MgSO₄, 1mM EGTA (ethylen-glycol-bis(2-aminoethylether)-N,N-tetraacetic acid) and 5% (v/v) paraformaldehyde. The paraformaldehyde stock solution (20% w/v) was dissolved by heating to 65°C, neutralised with NaOH and stored at -20°C. Once dissected, the tissues were transferred to sterile universal tubes containing 2 mls of fixing solution. An equal volume of methanol was added and the tissues were gently shaken for 5 minutes, and left to fix for 30 minutes - 2 hours. The fixing solution was removed and the tissues were transferred to eppendorf tubes containing 1 ml of 90% (v/v) methanol and 10% (v/v) 0.5M EGTA (ME). The ovaries were then refixed and dehydrated by passage through a series of steps consisting of ME and .PP (PP consists of 4% (v/v) paraformaldehyde in phosphate buffered saline (PBS, 130 mM NaCl, 10 mM sodium

phosphate pH 7.2)). The first step is for 5 minutes in 7/3 ME/PP, the second is for 5 minutes in 1/1 ME/PP, the third for 5 minutes in 3/7 ME/PP and the last step for 20 minutes in PP alone. The tissues were then washed in PBS for 10 minutes and directly subjected to the following steps. They can, if desired be stored at this point by dehydrating them in a 30%, 50% and 70%(w/v) ethanol series and leaving them at -20°C. Before proceeding with the next stage they must be rehydrated.

2.12.2 Probe Labelling Using Digoxigenin - Non Radioactive Method

Probes consisted of linearised plasmids containing cDNA inserts and were labelled according to the protocol supplied with the Beohringer Kit. The DNA to be labelled was denatured at 90°C for 10 minutes and rapidly cooled for 10 minutes on ice. Buffer containing unlabelled dATP, dGTP, dCTP, dGTP and digoxigenin labelled dUTP, the correct reaction salt conditions and primers, and Klenow were added and the mixture incubated overnight at 37°C. The labelling was done by random priming (Feinberg and Vogelstein, 1983). The probes were precipitated with an equal volume of 4 M Ammonium Acetate pH 6.6 and four volumes of absolute ethanol at -20°C overnight. 10µg of sonicated salmon sperm DNA was also added to the precipitation. The labelled DNA fragments and salmon sperm DNA were recovered by centrifugation in a microfuge for 5 minutes, washed in 70% (w/v) ethanol, air dried and resuspended in hybridisation solution (2.5.14.4).

2.12.3 Pretreatment of Tissues

All pretreatment steps were performed in 1.5 ml eppendorf tubes in 1 ml solution at room temperature on a revolving wheel. The tissues were first washed 3 times for 5 minutes each in PBT (PBS plus 0.1% (v/v) Tween 20). They were then incubated for 3-5 minutes in 50 µg/ml Proteinase K in PBS. The digestion was stopped by

incubating for 2 minutes in 2 mg/ml glycine in PBT. The tissues were then washed 2 times for 5 minutes each in PBT, refixed for 20 minutes with PP and finally washed 3 times for 10 minutes each in PBT.

2.12.4 Hybridisation and Washing

The hybridisation solution (HS) consists of 50% (v/v) formamide, 5X SSC, 50 µg/ml heparin, 0.1% (v/v) Tween 20 and 100 µg/ml sonicated and denatured salmon sperm DNA. (HS maybe stored at -20°C.) The tissues were washed for 20 minutes in 1/1 HS/PBT then for 20-60 minutes in HS. They were then transferred to a 45°C water bath for 20-60 minutes. Most of the supernatant was then carefully removed and the heat denatured probe (90°C, 5 minutes) was added (refer to 2.5.14.2). The probe concentration was aimed to be at approximately 0.5 µg/ml. The tissues were hybridised overnight at 42°C. The tissues were washed at room temperature for 20 minutes at each step. The first wash was in HS, the second in 4/1 HS/PBT, the third in 3/2 HS/PBT, the fourth in 2/3 HS/PBT, the fifth in 1/4 HS/PBT and the last two 20 minutes each in PBT.

2.12.5 Signal Detection

The antibody-conjugate solution (supplied with the Beohringer Kit) was used to detect the digoxigenin signal. The antibody was freshly diluted 1/2000 to 1/5000 in PBT and was preabsorbed for 1 hour with fixed tissues as this was found to be a critical step in reducing the background signal. The antibody solution (500µl) was then transferred to the experimental tissues and incubated for 1 hour at room temperature on a revolving wheel. The tissues were washed 4 times for 20 minutes each in PBT, then 3 times for 5 minutes each in 100 mM NaCl, 50 mM MgCl₂, 100 mM Tris pH 9.5, 1mM Levamisole (a potent inhibitor of lysosomal phosphatases), 0.1% (v/v) Tween 20. The tissues were then transferred to a microtitre dish with 250 µl of the above solution. The antibody signal was detected using a colour

reaction by adding 1.25 μ l of NBT (nitroblue tetrazolium salt, 75 mg/ml in dimethyl formamide, 70% (v/v)) and 0.9 μ l of the X-phosphate solution (5-bromo-4-chloro-3-indolyl phosphate, toluidinium salt, 50 mg/ml in dimethylformamide) both of which were supplied with the Boehringer kit. The colour developed after approximately 2 hours in the dark. The colour development was stopped by adding PBT before background colour appeared. The tissues were mounted in Gurr's Water Mounting Medium and flattened by placing a brass weight on top of the coverslip overnight. For longer term storage tissues were dehydrated by successive 15 minute incubations in absolute ethanol and absolute isopropanol (P. Lawrence, pers.comm.). Stored tissues were mounted in DPX mountant (BDH). The tissues were photographed at X10 magnification using Kodak Vericolor III professional film.

2.13 Detection of β -galactosidase expression patterns in situ

Flies or larvae were dissected in Ringers solution then transferred to the well of a microtitre dish containing 100 μ l of staining buffer [10 mM Na_2HPO_4 pH 7.0; 150 mM NaCl; 1 mM MgCl_2 ; 7.0 mM $\text{K}_4\text{Fe}(\text{CN})_6 \cdot 3\text{H}_2\text{O}$; 7 mM $\text{K}_3\text{Fe}(\text{CN})_6 \cdot 3\text{H}_2\text{O}$; 0.2% (w/v) 5-Bromo-4-chloro-3-indolyl-1-D-galacto-pyranoside (X-gal)]. Staining patterns were observed under the dissection microscope and were normally visible after 1 hour. Maximum staining normally occurred within 6 hours.

CHAPTER THREE

SELECTION AND PRELIMINARY MOLECULAR CHARACTERISATION OF TRANSPOSANT LINES EXHIBITING TESTIS - SPECIFIC β GALACTOSIDASE EXPRESSION

3.1 Screening Enhancer-trap lines for sex-specific β -galactosidase staining patterns

Multiple lines of flies carrying the enhancer trap transposon $p[lac Z, ry^+]$ were generated according to the crossing scheme illustrated in Fig 3.1A, by John Merriam (UCLA). A parental transformant, homozygous for an X chromosomal insertion of the transposon was crossed to a strain carrying a stable P-transposase source in the form of the "jumpstarter" element, $\Delta 2-3$ at chromosome position 99B (Robertson *et.al.*, 1988). Individual F1 male progeny from this cross, which carried both the enhancer trap and jumpstarter elements, were crossed to rosy females and male progeny were selected which carried novel insertions of $p[lac Z, ry^+]$, as a result of transposition of the element in their male parent. By this method, several hundred lines carrying single independent insertions were obtained. In a proportion of lines further crosses were carried out to yield progeny homozygous for the $p[lac Z, ry^+]$ insertion and identify its chromosomal location (Fig 3.1B). 232 lines, which will be referred to as transposants, were screened for sex specific adult β -galactosidase staining patterns, which potentially identified genes under sex-specific regulation. Table 3.1 summarises the staining patterns observed.

About 70% of the transposants exhibited some staining above background and in the vast majority of cases this included staining of gonadal and non gonadal tissue types. Where gonadal staining did occur it was much more likely to involve germline derived cells than somatic cells, in both males and females. There was also a significant preponderance of staining in the male gonad over staining in the female gonad. This may have been a reflection of preferential insertion of the mobilised transposon into transcriptionally active genes since transposition occurred in the male germline (discussed in Bownes 1990).

Table 3.1 Summary of adult β -galactosidase staining patterns of p[*lac Z*, *ry*⁺] transposants

Class	Number	%
Adults unstained	69	30
Male and female gonads and other tissue(s)	37	16
Tissues excluding gonads	19	8
Male and female gonads	50	22
Female gonad only	8	3
Male gonad only	49	21
Total	232	100

Figure 3.1A

Crossing scheme used to generate multiple $p[lac Z, ry^+]$ transposants

Females homozygous for an X chromosomal insertion of $p[lac Z, ry^+]$ were crossed to males carrying the Δ 2-3 "jumpstarter" element at 99B. ry^+ male progeny were selected, which carried novel $p[lac Z, ry^+]$ insertion on their 2nd or third chromosomes.

Figure 3.1B

Generation of homozygous $p[lac Z, ry^+]$ transposants

ry^+ flies from the cross illustrated in Fig 3.1A were crossed to a balancer strain carrying the T(2;3) CyO TM6 chromosome over marked second and third chromosomes (T(2;3). CyO TM6 balances chromosomes 2 and 3 simultaneously). ry^+ progeny with marked 2nd and 3rd chromosomes were backcrossed to the balancer strain. The genotypes of the progeny from the latter cross enabled determination of the chromosomal location (i.e. 2 or 3) of $p[lac Z, ry^+]$. Progeny of the indicated genotype were selected for the generation of homozygous lines.

A

$$\frac{P[lacZ ry^*]}{P[lacZ ry^*]}; \frac{ry^{506}}{ry^{506}} \times \frac{Sb e \Delta 2-3}{TM6}$$

$$\downarrow$$

$$\frac{ry^{506}}{ry^{506}} \times \frac{P[lacZ ry^*]}{ry^{506}}; \frac{ry^{506}}{Sb e \Delta 2-3}$$

Transposase from $\Delta 2-3$ causes $P[lacZ ry^*]$ to move

$$\frac{ry^{506}}{ry^{506}} \times ry^+ \text{ males - these will contain a } P[lacZ ry^*] \text{ in their genome}$$

$$\downarrow$$

Line with $P[lacZ]$ insertion

$$\downarrow$$

Stain for β -galactosidase to find pattern of expression

B

$$\frac{T(2;3) CyO TM6}{pr cn; mwh ry e} \times \frac{ry^{506}}{ry^{506}} [ry^*] \text{ from line}$$

$$\downarrow$$

$$\frac{T(2;3) CyO TM6}{pr cn; mwh ry e} \times \frac{+}{pr cn}; \frac{ry^{506}}{ry e} \quad P[lacZ ry^*] \text{ at one or other chromosome}$$

$$\downarrow$$

Select from progeny

$$\frac{T(2;3) CyO TM6}{+; ry^{506}} \times \frac{CyO TM6}{+; ry^{506}}$$

with $P[lacZ ry^*]$ on either chromosome to establish line that will generate homozygotes

The only examples of sex specific staining were observed in gonadal tissue. A decision was taken to further characterise a number of transposants exhibiting testis specific staining. This choice was made in recognition of the fact that in comparison with oogenesis, there is relatively little information regarding the molecular basis of spermatogenesis. As discussed in the introduction, a considerable number of genes involved in the execution of oogenesis and the related process of early embryogenesis have been cloned and characterised in detail. In contrast, the literature contains few examples of genes involved in spermatogenesis which have been characterised to this extent. Furthermore, as discussed in chapter one, the limited molecular analysis which has been undertaken so far has yielded valuable insights into the translational control which is so essential to the process, as well as mechanisms of transcriptional regulation (Michiels *et.al.*, 1989; Kuhn *et.al.*, 1988; Schäfer *et.al.*, 1990). This information has opened the way to identifying some of the regulatory genes involved and as additional structural genes are cloned it should be possible to determine the degree of generality of the regulatory mechanisms characterised thus far and to identify related systems.

In the testis, a variety of β -galactosidase staining patterns were observed, involving both somatic and germline derived cell types. Most testis staining patterns involved the spermatogenic cysts; either germline derived cells, somatic cyst cells or both. However, there was variation between transposants in the stages of spermatogenesis at which staining was observed. For example, some lines stained exclusively at early spermatogonial stages in the apical region of the testis whereas others stained throughout the major part of the spermatogenic programme or exclusively at later stages of the differentiation process. Staining outwith the spermatogenic cysts was observed in relatively few transposants and typically involved the pigment cells of the testis sheath or terminal epithelial

cells at the base of the testis.

3.1.1 Preliminary molecular characterisation of "testis-specific" lines.

As a first step towards the isolation of the endogenous genes identified by the testis-specific staining patterns, it was necessary to clone genomic sequences flanking the p[*lac Z ry*⁺] insertion site. This required the construction of a genomic library from the transposant of interest, which could then be screened with labelled p[*lac Z ry*⁺] sequences. The strategy adopted was dictated by the availability of restriction sites within p[*lac Z ry*⁺] which were compatible with the cloning site of a suitable lambda vector. As can be seen in Fig 1.3, there is a Sal I site at the 3' end of the *lac Z* sequences in p[*lac Z ry*⁺]. Sal I digestion of genomic DNA from a transposant generates restriction fragments consisting of the 5' end of p[*lac Z ry*⁺], including the *lac Z* sequence, plus a variable amount of genomic DNA, depending on the distance to the nearest endogenous Sal I site. In order to maximise the probability of obtaining sequences of interest from the library it was desirable to clone as large a Sal I fragment as possible, subject to the restrictions imposed by the cloning capacity of the available vectors. The lambda replacement vector EMBL 4 appeared suitable for the purpose since it accepts Sal I fragments in the range 10-22 kb.

Southern hybridisation was employed to determine the size of the desired Sal I fragments in seven transposant lines and to confirm that each line contained a p(*lac-Z,ry*⁺) insertion at only a single chromosomal site. These lines, all of which had been made homozygous for the transposon, were selected because they exhibited a representative range of β -galactosidase staining patterns, exclusive to the testes. Genomic DNA was extracted from each line, digested to completion with Sal I, fractionated on an agarose gel and Southern blotted. The filters were probed with a

radioactively labelled Pst I fragment of the *E.coli lac Z* gene excised from the plasmid pMC 1871 (see materials). The Sal I fragments recognised by the *lac Z* probe ranged in size from 4-13 kb (Fig 3.2A and 3.2B). The staining patterns in each of the seven lines, together with the corresponding Sal I fragment sizes and the chromosomal locations of p[*lac-Z,ry*⁺] are presented in Table 3.2. Transposant lines K8 and A22 which stained in the spermatogenic cysts were selected for further analysis on two criteria. Firstly, both these transposants exhibited staining of a relatively high intensity throughout a major portion of the spermatogenic programme. It was therefore reasoned that the staining patterns in these transposants might identify endogenous genes which encoded relatively abundant transcripts, which would ease Northern analysis of the genes' expression patterns. Secondly, the largest Sal I fragments were detected in these two transposants and both fragments were within the acceptable size range (approximately 13 and 11 kb respectively). The patterns of β -galactosidase expression observed in these two lines are shown in Fig 3.3B and D.

3.2 Screening K8 and A22 genomic libraries

K8 and A22 Sal I genomic libraries were constructed in EMBL 4 and screened with the *lac Z* probe. Single positive clones, K8 E4 and A22 E4, were isolated from the respective libraries and DNA was prepared from each clone and digested with Sal I. These digests were run on an agarose gel in parallel with Sal I digested DNA from the corresponding transposant, Southern blotted and probed with *lac Z* to confirm the identity of the cloned fragments (Fig 3.2C). In both cases, it appeared that the desired fragment had been cloned since it was recognised by the *lac Z* probe and co-migrated with the fragment detected in genomic DNA. When Southern blots of Sal I digested genomic DNA from transposants K8 and A22 were probed with the

cloned fragments from the corresponding lines only single bands of hybridisation were observed, indicating that both fragments contained single copy *Drosophila* sequences (not shown).

3.3 Identifying the chromosomal locations of the insertions

Salivary gland polytene chromosome squashes were prepared from the two transposants and probed with the Pst I *lac Z* fragment from plasmid pMC 1871 to determine the chromosomal positions of the p[*lac Z ry*⁺] insertions. Single insertion sites were detected in both K8 and A22, at 3R 99B and 1 8C-9A respectively (Fig 3.3A and C). The K8 insertion site is close to the position of a previously identified gene *mst 98CE*, which is transcribed exclusively in the male germline (Di Benidetto *et.al.*, 1987) suggesting that cis-acting sequences associated with this gene might be responsible for the β -galactosidase staining pattern in this transposant. The A22 insertion site does not correspond to the position of any reported gene with testis expression or male sterile mutation, increasing the likelihood that the testis-specific staining pattern observed in this transposant identified a novel gene with a role in spermatogenesis, situated next to the transposon insertion site. With the objective of isolating the endogenous genes identified by the staining patterns in K8 and A22, a further molecular analysis was undertaken of the genomic regions into which the transposon had inserted in these two lines.

3.4 Location of transcription units in the vicinity of the K8 and A22 insertions

A Northern analysis was conducted in order to determine whether the cloned genomic regions from transposants A22 and K8 contained any coding sequences with

a testis specific expression pattern. Sal I insert fragments were purified from the genomic clones A22 E4 and K8 E4 and used to probe Northern blots of total RNA extracted from male and female OrR flies. Neither the A22 or K8 Sal I fragments detected any clear transcripts in the Northern hybridisation and there was certainly no evidence of a male-specific transcript in either case (not shown). Two approaches were taken to overcome possible limitations of the above strategy. Firstly, since it was possible that the genes of interest lay outwith the Sal I fragments, chromosome walks were undertaken to extend the cloned genomic regions available for analysis. The second approach was designed to improve the sensitivity of the Northern analysis. This had previously been limited by the large size of the Sal I probe fragments. It was possible that any transcribed sequences formed only a small fraction of the total Sal I fragment, in which case the probe would have low specificity and sensitivity for transcript detection. To overcome this problem, a "reverse Northern" approach was adopted in order to identify transcribed sub-fragments within the genomic clones. This involved digestion of the genomic clones from the two transposants with suitable restriction enzymes, followed by bidirectional Southern blotting. The duplicate filters were then probed with ³²P dCTP labelled male or female cDNA, prepared by reverse transcription of poly (A)⁺ RNA extracted from whole flies. Transcribed fragments identified in this way could then be used as probes in conventional Northern hybridisation.

3.4.1 Transposant K8

The K8 Sal I fragment was used to screen a random genomic library consisting of overlapping Sau 3A1 fragments cloned into the Bam H1 site of the vector lambda DASH. Two genomic clones were identified in this screen: K8 LD1 and K8 LD2. A crude restriction map was derived of the entire cloned genomic region represented in the three lambda

Figure 3.2

A and B: Southern blots of Sal I digested genomic DNA from p[*lac Z*, *ry*⁺] transposant lines exhibiting testis-specific β -galactosidase expression, probed with *lac Z*

A: 1 = line D7
 2 = line D13
 3 = line J5
 4 = line K8

B: 1 = line A22
 2 = line A18
 3 = line A12

C: Southern blot of Sal I digested genomic DNA from transposants A22 and K8 electrophoresed in parallel with the cloned Sal I fragments from the corresponding lines and probed with *lac Z*

 1 = A22 Sal I insert
 2 = A22 genomic DNA
 3 = K8 genomic DNA
 4 = K8 Sal I insert

D: Reverse Northern analysis of the cloned genomic regions surrounding the p[*lac Z*, *ry*⁺] insertion sites in transposants A22 and K8

Genomic clones double digested with the indicated restriction enzymes were southern blotted and probed with radiolabelled cDNA made from male poly(A)⁺ RNA. An indistinguishable result was obtained when a duplicate blot was probed with radiolabelled cDNA prepared from female flies.

 1 = A22 LD3 X Eco RI \ Sal I
 2 = A22 E4 X Eco RI \ Pst I
 3 = K8 LD2 X Eco RI \ Sal I
 4 = K8 LD1 X Eco RI \ Sal I
 5 = K8 E4 X Eco RI \ Pst I

The indicated restriction fragments of A22 LD3 and K8 LD1 were isolated for further analysis.

Figure 3.6A

Northern blot of total RNA from isolated OrR tissues probed with the genomic fragment A-5.1

Approximately 10 µg of RNA were loaded per lane.

- 1 = testis
- 2 = male carcass
- 3 = ovary
- 4 = female carcass

1.0 kb and 1.6 kb ovary enriched transcripts and an apparently ovary specific 2.5 kb transcript are indicated.

Figure 3.6B

Southern hybridisation to genomic clones A22 E4 and A22 LD3 with cDNA OV.3

- 1 = A22 LD3 X Eco RI \ Sal I
- 2 = A22 E4 X Eco RI

The 3.5 kb Eco RI \ Sal I fragment of A22 LD3 and the 3.5 kb Eco RI fragment of A22 E4 to which OV.3 hybridises are indicated

Figure 3.6C

Southern hybridisation to genomic clones A22 E4 and

A22 LD3 with cDNA OV.1

- 1 = A22 E4 X Eco RI \ Bam HI
- 2 = A22 LD3 X Eco RI \ Sal I

The 1.6 kb Eco RI \ Sal I fragment of A22 LD3 to which OV.1 hybridises is indicated. Weak hybridisation is also apparent to the 3.5 kb Sal I fragment of A22 LD3 and the 3.1 kb Eco RI \ Bam HI fragment of A22 E4.

Figure 4.1A

Northern blot of RNA from isolated OrR tissues probed with cDNA OV.1

Approximately 10 µg of RNA were loaded per lane

- 1 = male carcass
- 2 = testis
- 3 = female carcass
- 4 = ovary

The 1.0 kb, 1.6 kb and 2.5 kb transcripts indicated probably correspond to the transcripts detected by A-5.1 (Fig 3.6A). A 1.0 kb ovary enriched transcript and a 0.8 kb testis-specific transcript are also indicated.

Figure 4.1B

Northern blot of RNA from isolated OrR tissues probed with cDNA T1

Approximately 10 µg of RNA were loaded per lane

- 1 = male carcass
- 2 = testis
- 3 = female carcass
- 4 = ovary

The 0.8 kb, 1.6 kb, 1.9 kb and 2.5 kb transcripts detected probably correspond to transcripts detected by cDNA OV.1 (Fig 4.1A).

Figure 4.1C

Southern blot of Eco RI digested OrR DNA probed with cDNA T1

The single signal corresponds to a fragment size of approximately 7.5 kb.

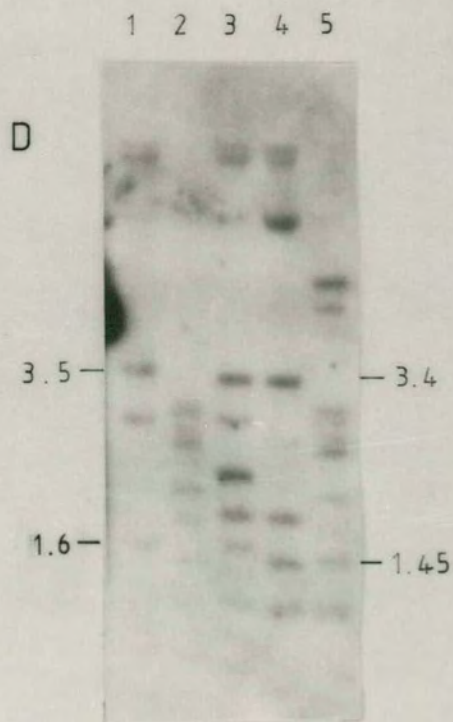
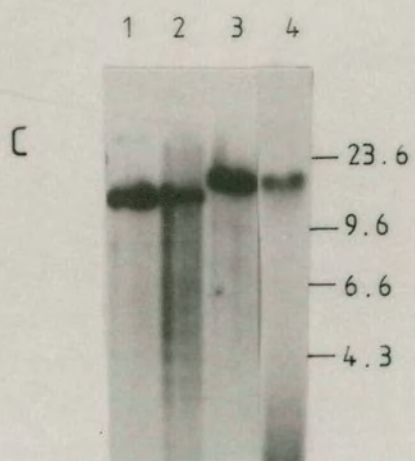
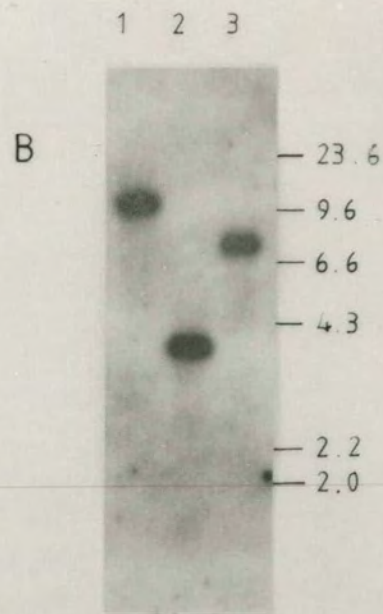
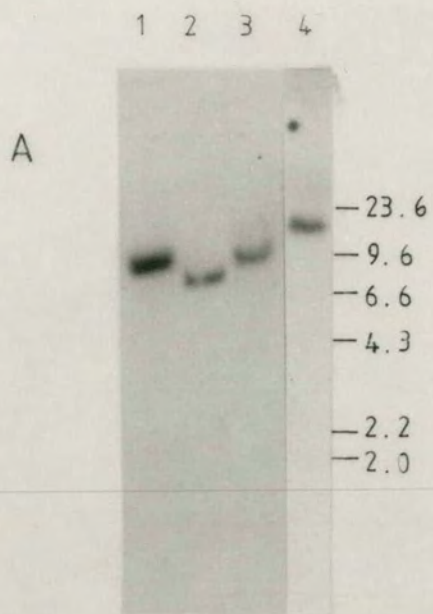


Table 3.2 β -galactosidase staining patterns, sizes of 5' flanking Sal 1 fragments and chromosomal locations of p[lac-Z, ry⁺] insertions in seven transposants

Transposant Line	Staining Pattern	Sal 1 fragment size (kb)	Chromosomal location (where determined)
A12	spots of staining over entire testis sheath-probably individual pigment cells	7.5	2
A18	staining of cyst cells at all stages of spermatogenesis	4.0	2
A22	germline-specific staining from spermatogonial stage in apical region to spermatid stage. Staining in spermatogonial and spermatocyte stages of larval gonad. ¹	10.7	2
D7	indistinguishable from A22	9.4	3
D13	germline-specific staining in spermatogonial stage and perhaps also stem cells	7.5	3
J5	germline staining from stem cell or spermatogonial stage until meiosis	9.5	2
K8	indistinguishable from A22 or D7	12.7	3

¹ Staining of larval gonads was only investigated in this transposant.

Figure 3.3

A and C : *in situ* hybridisation to salivary gland polytene chromosome squashes of transposant K8 and A22 with tritiated *lac Z* probe

Magnification = X 100

A = K8

C = A22

The hybridisation signal is indicated at 8C-9A. The landmark X chromosome puff at 2B is also indicated.

E: *in situ* hybridisation to salivary gland polytene chromosome squash of transposant A22 with digoxigenin labelled OV.1 cDNA

Magnification = X 100

B and D: β -galactosidase staining patterns in adult testes of transposants K8 and A22

TER = terminal end of testis

API = apical end of testis

C = spermatogenic cyst

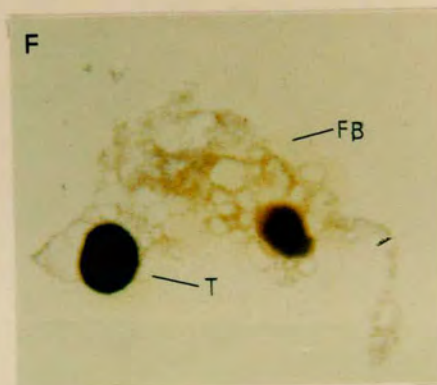
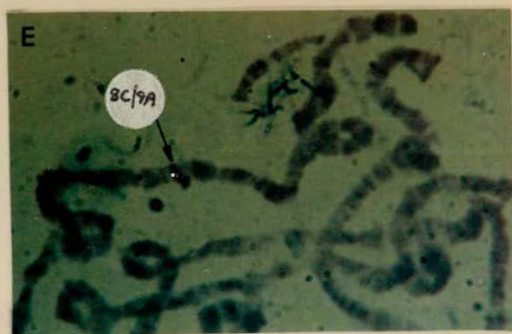
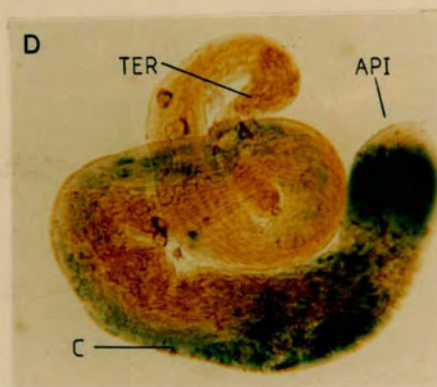
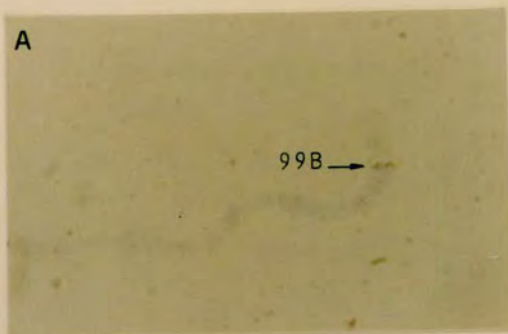
Both lines stain in the spermatogenic cysts throughout a large proportion of the spermatogenic programme from the spermatogonial stage at the apex up to post meiotic stages beyond the testis midpoint. Maximum intensity is seen at the apical end of the testis in the region of the mitotic gonial proliferation.

F: β -galactosidase staining in third instar larval gonads of transposant A22

T = testis

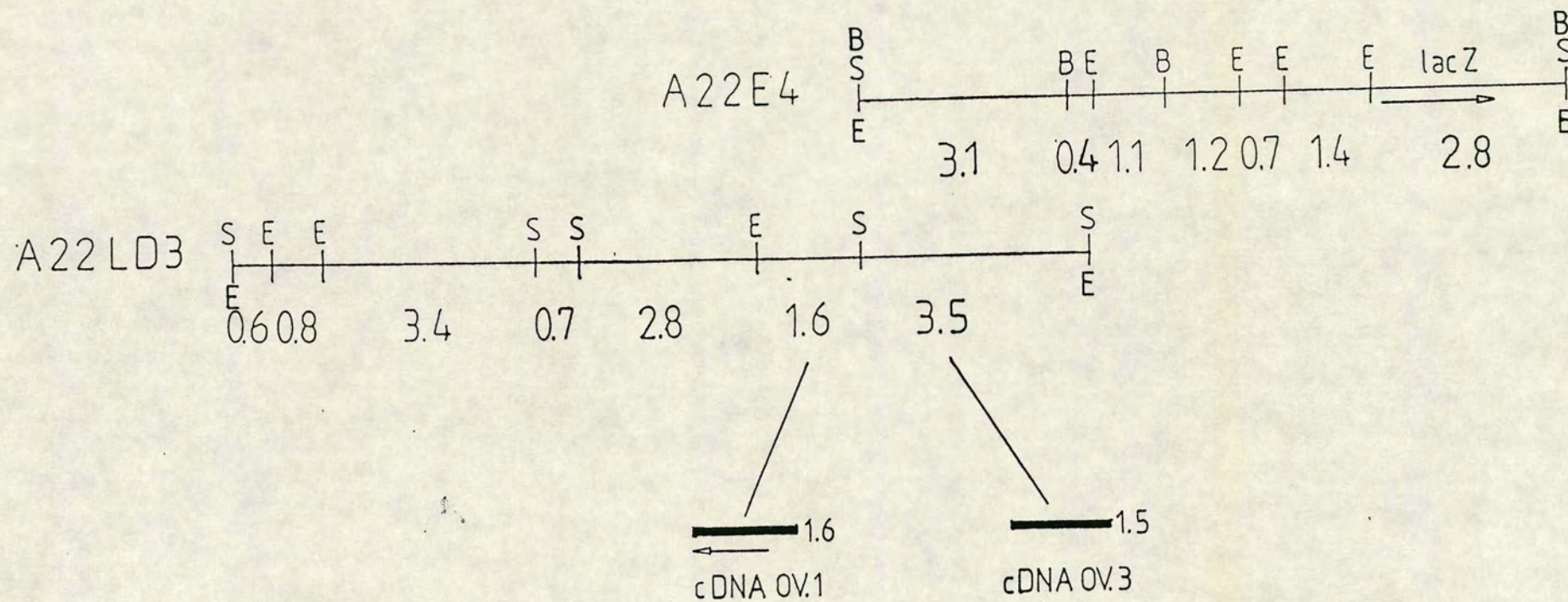
FB = larval fat body

Intense staining is seen over the whole testis and is absent from the surrounding fat body. The uniform staining over the larval gonad is consistent with β -galactosidase activity in both the spermatogonial and spermatocyte stages which are present in these immature testes (E. Kübli, pers. comm.)



Correction

Subsequent examination of transposant line A22 has revealed that it is an heterogeneous stock carrying two p[lac Z,ry⁺] insertions at 38C and 8C-9A. The IMPD gene actually maps to the vicinity of the insertion at 8C-9A.



overlaps with cDNA T1 which maps outside
this genomic region, probably in the
upstream direction from p[lac-Z, ry⁺].

clones and patterns of cross hybridisation between the clones were used to confirm regions of overlap (Fig 3.4). The results indicated that the sequences in K8 LD1 and K8 LD2 extended the cloned region 10.4 kb upstream of the transposon insertion site and 3.45 kb downstream.

The three K8 genomic clones (K8 LD1, K8 LD2 and K8 E4) were double digested with a combination of Eco RI and Sal I or Pst I and subjected to reverse Northern analysis (Fig 3.2D). None of the sub-fragments hybridised exclusively with male or female cDNA but a number of sub-fragments bound cDNA from both sexes. A 3.4 kb Eco RI / Sal I fragment of K8 LD1 upstream of p[lac Z, ry⁺] (Fig 3.4), which gave a relatively intense hybridisation signal (Fig 3.2D) was gel purified and used to probe a Northern blot of OrR RNA. To further improve the efficiency of transcript detection, the RNA used was prepared from isolated male and female gonads and non-gonadal tissues (carcasses). However, once again no clear transcripts were detected with this probe in the RNA of any of these tissues (Data not shown). When the same procedure was followed with the 1.45 kb Sal I fragment of K8 LD1 downstream of the transposon a negative result was also obtained and I turned to a similar investigation of the other transposant line, A22.

3.4.2 Transposant A22

The A22 Sal I fragment identified a single overlapping clone in the lambda DASH library (A22 LD1), which extended the cloned genomic region by 9 kb upstream of the transposon (Fig 3.5). A 5.1 kb Eco RI / Sal I fragment (A-5.1) of A22 E4 which appeared to contain sequences which were transcribed in males and females, on the basis of the reverse Northern result (Fig 3.2D), was selected for gel purification. When A-5.1 was used to probe a Northern blot of RNA from dissected *Drosophila* tissues it detected three transcripts of approximately 1.0 kb, 1.6 kb and 2.5 kb

(Fig 3.6A) Most abundant is the 1.0 kb transcript which appeared highly ovary enriched, with only faint signals detectable in males or female carcasses. The 1.6 kb and 2.5 kb transcripts are of similar, relatively low abundance but whilst the larger transcript was only detected in ovary RNA (lane 3), the 1.6 kb transcript, although highly ovary enriched, also appeared to be present at a low level in the other tissues examined. Additional transcripts could have been masked by the evident non-specific hybridisation to the highly abundant rRNA.

Figure 3.5

Restriction maps of two overlapping clones from the genomic region upstream of the p[*lac Z*, *ry*] insertion site in transposant A22

E = Eco RI

S = Sal I

B = Bam HI

Fragment sizes are in kb. Arrows indicate direction of transcription.

A22 LD3 is a Sau 3AI partial digestion product. The terminal Sal I sites of A22 LD3 which are indicated are within the vector polylinker.

A22 E4 is the original clone isolated from the A22 Sal I library and includes *lac Z* sequences from p[*lac Z*, *ry*⁺]. A22 LD3 is an overlapping clone isolated from the Canton S random genomic library.

The genomic fragments to which cDNA clones OV.1 and OV.3 were mapped by Southern hybridisation are indicated. Together these two fragments comprise the 5.1 kb Eco RI \ Sal I fragment A-5.1.

cDNA OV.1 overlaps with cDNA T1 but cDNA T1 does not hybridise anywhere within this genomic region.

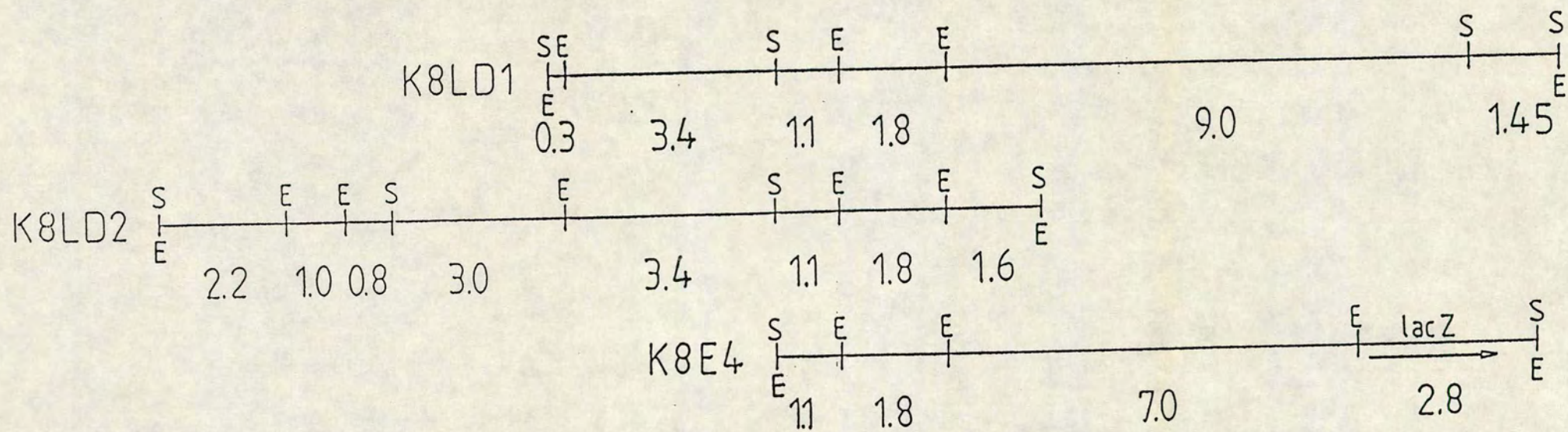


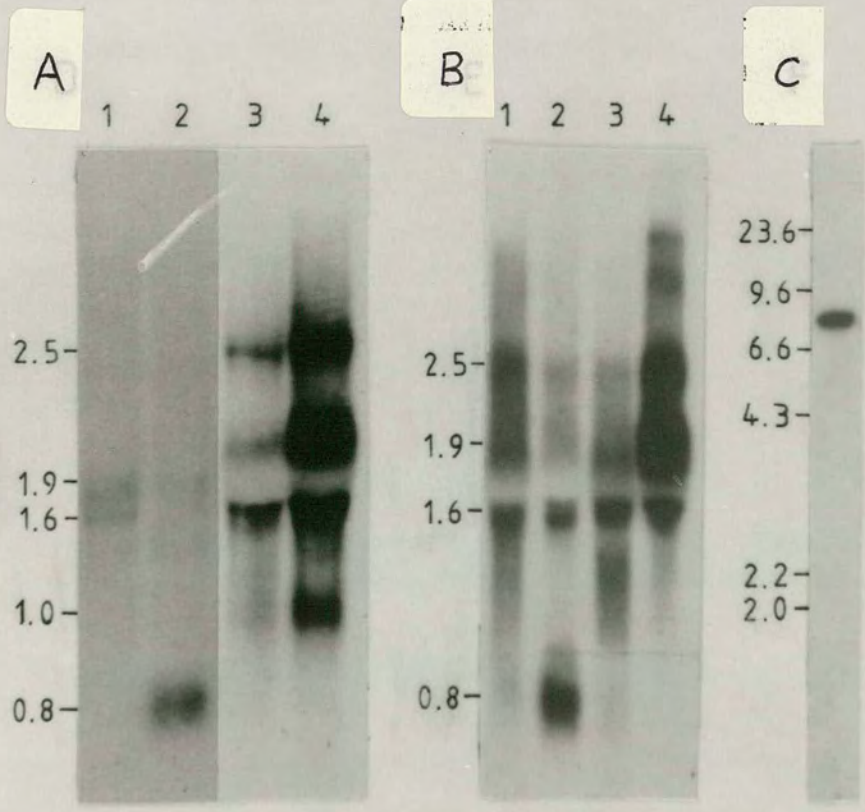
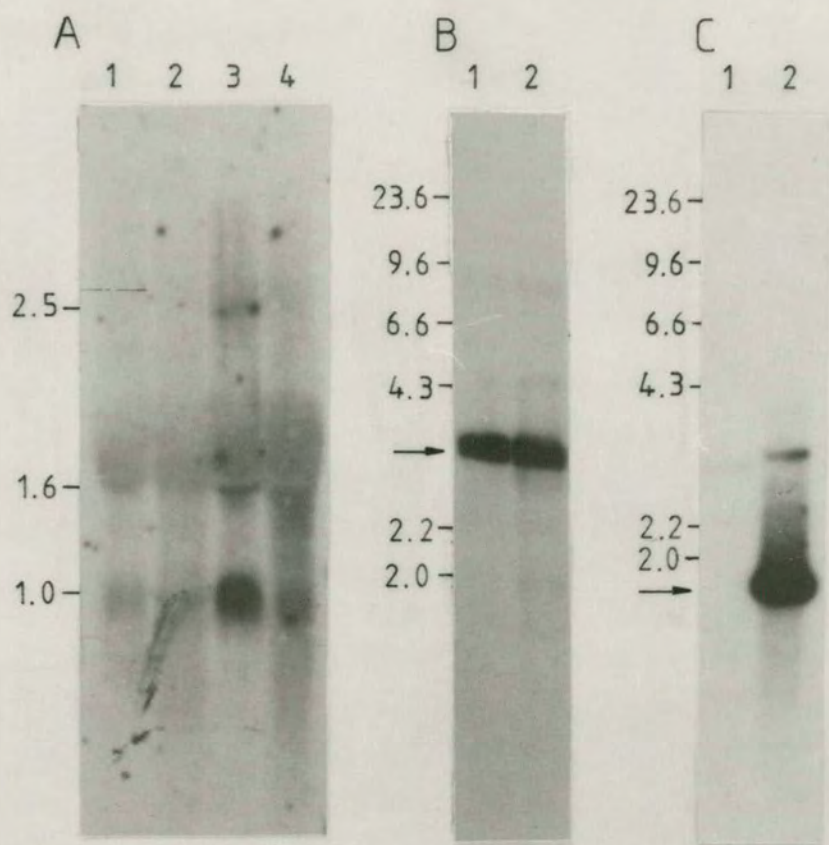
Figure 3.4

Restriction maps of three overlapping clones from the genomic region around the p[*lac Z*, *ry*⁺] insertion site in transposant K8

E = Eco RI

S = Sal I

Fragment sizes are in kb. K8 LD1 and K8 LD2 are Sau 3AI partial digestion products. The indicated terminal Sal I sites of these clones are within the vector polylinker. K8 E4 is the original clone isolated from the K8 Sal I genomic library with a *lac Z* probe and includes *lac Z* sequences from p[*lac Z*, *ry*⁺]. K8 LD1 and K8 LD2 are overlapping clones isolated from a random Canton S genomic library.



CHAPTER FOUR

MOLECULAR CHARACTERISATION OF A GENE IN THE VICINITY OF THE p[*lac Z*,ry⁺] INSERTION IN A22

4.1 Isolation of cDNA clones representing transcribed sequences within A-5.1

The Northern experiments described indicated that the genomic fragment A-5.1 contained sequences which were transcribed at a high level in adult ovaries. The existence of a gene with this pattern of expression, in the vicinity of p[lac Z ry⁺] had not been predicted by the testis-specific β -galactosidase staining pattern observed in this line. It therefore seemed unlikely that the transcribed sequences within A-5.1 belonged to a gene identified by this staining pattern. However, the data did not exclude the possibility that the gene giving rise to the 1.0 kb and 1.6 kb transcripts extended beyond A-5.1 and included sequences which were transcribed in a testis specific fashion. Consequently, it was decided to proceed further with the investigation of this transcription unit.

In order to begin to understand the organisation of the coding sequences within A-5.1 the first objective was the isolation of cDNA clones representing the transcripts recognised by A-5.1 on a Northern blot (Fig 3.6A). When an adult ovary cDNA library was probed with this fragment, three positive clones were isolated. Two of these clones OV.1 and OV.2 were evidently duplicate isolates since they contained inserts of 1.6 kb which appeared to be identical, on the basis of restriction patterns and cross hybridisation data. The third clone, OV.3, contained an insert of 1.5 kb which was evidently unrelated in sequence to OV.1 and OV.2.

4.1.1 Mapping of cDNA clones OV.1, and OV.3 on to the cloned genomic region

cDNA's OV.1 and OV.3 were assigned to positions on the genomic restriction map by Southern analysis. Restriction digests of A22 E4 and A22 LD3 were Southern blotted and probed with labelled cDNA's. OV.1 hybridised strongly to

the 1.6 kb Eco RI \ Sal I sub-fragment of A-5.1 and weakly to the adjacent 3.5 kb Sal I sub-fragment of A-5.1 and the 3.1 kb Eco RI \ Bam HI fragment of A22 LD3 (Fig 3.5 and Fig 3.6C). OV.3 hybridised to the 3.5 kb Sal I sub-fragment of A-5.1 and the 3.5 kb Eco RI fragment of A22 LD3 (Fig 3.5 and Fig 3.6B). The conclusions from these results are as follows: 1) The genomic sequences within the cloned region depicted in Fig 3.5, which are represented in cDNA OV.1, are largely contained within the 1.6 kb sub-fragment of A-5.1 but extend into the adjacent 3.5 kb Sal I fragment. 2) The sequences represented by cDNA OV.3 are confined to the 3.5 kb Sal I fragment of A22 LD3.

4.2 A-5.1 includes part of a gene encoding multiple independently regulated transcripts

OV.1 detected multiple transcripts with a variety of expression patterns on a Northern blot (Fig 4.1A). These include three transcripts of 1.0 kb, 1.6 kb and 2.5 kb which probably correspond to the transcripts detected by A-5.1 (Fig 3.6A), although differences do exist between the expression profiles revealed by the two probes. Specifically, the 2.5 kb transcript which appeared ovary specific in Fig 3.6A is clearly detectable in the female carcass with the OV.1 probe (Fig 4.1A lane 3) and the 1.0 kb transcript which was detected in all tissues with A-5.1 appears to be female-specific in Fig 4.1A. However, it seems possible that these discrepancies are artefactual. Firstly, the level of the 2.5 kb transcript may have been below the detectable threshold in tissues other than the ovary in the earlier experiment (Fig 3.6A) and might have appeared after longer autoradiographic exposure. Secondly, on the basis of ethidium bromide staining of the RNA gel represented in Fig 4.1A, lanes 3 and 4 appeared to be slightly overloaded relative to lanes 1 and 2 which could account for the failure to detect the 1.0 kb transcript in the male tissues with the OV.1 probe.

In addition to the above three transcripts, OV.1 recognises a testis-specific transcript of 0.8 kb (Fig 4.1 lane 2) and an ovary enriched transcript of 1.9 kb which is present at a reduced level in the other tissues. Neither of these transcripts was detected by A-5.1 (Fig 3.6) although a 1.9 kb transcript could have been masked by the ribosomal smear.

OV.3 did not detect any of the above transcripts and actually failed to detect any discernible transcripts at all (not shown). This finding, combined with the observations that OV.3 does not share any sequence homology with OV.1 on the basis of cross hybridisation and recognises a different genomic sub-fragment (Fig 3.5) suggests that OV.3 represents a transcript from a separate but closely neighbouring gene. This transcript is presumably too rare to detect in a Northern blot of total RNA or was masked by non-specific hybridisation to rRNA.

The above data indicated that OV.1 shares sequence homology with a number of transcripts including those originally detected by the genomic fragment A-5.1 and since the A22 E4 genomic clone appeared on the basis of Southern hybridisation, to contain only single copy sequences, it seemed most likely that these transcripts were overlapping products of a single gene (gene A). It was also concluded from this Northern analysis that gene A must extend beyond the genomic region represented in the A22 E4 and A22 LD3 clones. The evidence for this was twofold:

- 1) OV.1 detects transcripts which are not detected by the genomic fragment A-5.1, including a 0.8 kb testis specific species and a 1.9 kb ovary enriched species (Fig 4.1A).

- 2) OV.1 does not recognise any sequences outside A-5.1, within the cloned region.

It appears that a separate gene (gene B), giving rise to the transcript represented by OV.3, lies immediately to the right of gene A, according to the restriction map in Fig 3.5. Therefore, it must be assumed that the proposed

outlying sequences of gene A are situated upstream of the p[lac Z, ry⁺] insertion and it follows that this gene must include an intron or introns totalling at least 8.3 kb (Fig 3.5).

The Northern data constituted evidence for the existence of a gene in the vicinity of the A22 p[lac Z, ry⁺] insertion which generates a number of transcripts, one of which is expressed exclusively in the testis. These observations raised the possibility that regulatory sequences associated with this gene could be responsible for the testis-specific β -galactosidase staining in this transposant. It could be envisaged for example, that it is a testis specific enhancer element flanking this gene which directs the expression of the 0.8 kb transcript and that this enhancer is acting upon the P-element promoter in this transposant. This raises the question of where, in relation to the transposon, this putative enhancer lies and why the β -galactosidase expression pattern does not mirror all aspects of the expression pattern of this gene, including the ovarian expression, if they are also governed by cis-acting elements. It is of course conceivable that the organisation of the various enhancer elements is such that only the one directing the testis-specific transcript is situated close enough to the 5' end of the transposon to influence the P-element promoter. However, the fact that a another transcription unit evidently separates this gene from p[lac-Z, ry⁺] would seem to exclude this possibility, unless the putative testis-specific enhancer element is acting across the intervening gene upon the P-element promoter, a model for which the author is aware of no published precedent. In considering these questions, it must be borne in mind that the multiple transcripts arising from this gene could result from differential promoter usage, pre-mRNA splicing, polyadenylation or any combination thereof and may not even require tissue-specific enhancer elements. Clearly, one can only speculate as to the mechanisms involved until a much more

thorough molecular analysis of this genomic region has been conducted.

4.3 Isolation and characterisation of a cDNA clone representing the testis specific transcript

The results discussed in the preceding sections indicated that OV.1 represents a 1.6 kb transcript which overlaps with a 0.8 kb testis specific transcript from the same gene. In order to determine the structural relationship between these two transcripts it was necessary to obtain a cDNA clone representing the 0.8 kb transcript. When OV.1 was used to screen an adult male cDNA library (section 2.3.5.3), two positive clones (T1 and T2) were isolated and since both contained apparently identical inserts of about 0.8 kb it seemed likely that they represented the testis specific transcript detected in the Northern analysis. On a genomic Southern blot T1 recognises a single Eco RI fragment of around 7.5 kb (Fig 4.1C) and on a Northern blot it detects all the same transcripts as OV.1, apart from the 1.0 kb ovary enriched transcript (Fig 4.1B). However, in this case the 2.5 kb transcript is clearly present in all tissues examined although highly ovary enriched. Taken together, the above findings with the T1 probe confirmed earlier evidence that this chromosomal region contains a gene which gives rise to multiple overlapping transcripts, including testis-specific and ovary enriched variants.

Southern analysis revealed that T1 does not hybridise anywhere within the cloned A22 genomic region. This supported the earlier evidence that the transcription unit must extend beyond the cloned region and includes an intron(s) of at least 8.3 kb separating fragment A-5.1 from the sequences encoding the testis-specific transcript.

4.4 Isolation of genomic DNA encoding the testis-specific transcript

Having established that the entire transcription unit was not contained within the existing genomic clones (A22 E4 and A22 LD3), an attempt was made to obtain the missing DNA. First, sequences corresponding to the 0.8 kb cDNA were isolated from the random genomic library by screening with a purified T1 insert. A total of three distinct clones (A22 LD4-A22 LD6) were obtained, with inserts ranging from 13.9-18 kb (Table 3.3). However, the genomic sequences contained in this set of clones were evidently not continuous with the previously isolated region, since none of the clones in the second set hybridised with A22 E4 or A22 LD3 in a Southern blot (not shown). In an effort to obtain the intervening DNA, the random genomic library was rescreened for additional overlapping clones, using A22 LD6 as a probe. In this second screen, to impose a degree of selection for sequences which would extend the existing cloned region, "end-specific" RNA probes were used which were generated from the T7 or SP6 promoters flanking the lambda DASH cloning site. A further pair of novel clones with inserts of 15.2 kb and 16.5 kb (A22 LD7 and A22 LD8; Table 3.3) was obtained with the T7 and SP6 generated probes respectively, but again neither clone was found to overlap with the A22 genomic region represented in Fig 3.5. The failure to obtain a continuous set of clones from the genomic library implied that the extent of the proposed intronic region might be significantly greater than the minimum value of 8.3 kb. However, since restriction maps were not derived for the later sets of clones (A22 LD4-A22 LD8) it is not possible to make a new estimate of the minimum distance in the genome which separates the sequences within A-5.1 which are represented in the OV.1 cDNA, from the sequences encoding the testis specific transcript. In subsequent sections, other possible reasons are discussed for the inability to link the genomic clones.

Table 3.3 Clones Isolated from OrR Genomic Library

λ Clone	Insert size	Component Eco R1 fragments	Eco R1 fragment recognised by T1 cDNA probe
A22 LD4 ^a	13.9	2.4, 4.5, 7.0	4.5
A22 LD5 ^a	15	9.0, 6.0	6.0
A22 LD6 ^a	18	8.0, 7.5, 2.5	7.5
A22 LD7 ^b	15.2	8.0, 5.0, 2.2	5.0
A22 LD8 ^c	16.5	8.0, 8.5	/

^aIsolated with T1 cDNA probe. ^bIsolated with "end-specific" RNA probe generated from A22 LD6 T7 promoter. ^cIsolated with RNA probe generated from A22 LD6 SP6 promoter.

CHAPTER FIVE

SEQUENCE ANALYSIS OF T1 and OV.1

5.1 Sequences of T1 and OV.1

T1 and OV.1 were sequenced, according to the strategy illustrated in Fig 5.1, in order to clarify the structural relationship between the two cDNA's and determine their coding capacities. The nucleotide sequences and predicted amino-acid sequences of T1 and OV.1 are presented individually in Figs 5.2 and 5.3 and the relationship between the two cDNA's and their coding regions is shown in Fig 5.1. This relationship is illustrated diagrammatically in Fig 5.4.

5.1.1 T1 cDNA

T1 is 844 bp in length and contains a long open reading frame (ORF) of 205 amino acids extending from the 5' terminus to nt. 615 (Fig 5.2). There is no ATG start codon, indicating that T1 is probably a 5' truncated copy of the mRNA and the absence of a polyadenylation signal or poly (A) tail from the 3' end of the cDNA suggests that it is also 3' truncated. This evidence that T1 is only a partial copy of an mRNA reduces the probability that this cDNA represents the testis-specific transcript detected by the OV.1 probe, since the transcript appears to be only around 0.8 kb in length (Fig 4.1A and B). However, this remains possible if only short sequences are missing at each end of cDNA T1.

5.1.2 OV.1 cDNA

OV.1 is 1613 bp in length excluding a poly (A) tail of 22 nucleotides and nucleotides 623 to 1226 of this cDNA are identical to nucleotides 241 to 844 of T1. (Fig 5.1). The extent of the OV.1 ORF could not be determined with certainty due to some remaining ambiguities in the 5' portion of the sequence. The provisional sequence as presented in Fig 5.3 contains an ORF of 190 amino acids,

Figure 5.1

Alignment of the nucleotide sequences of cDNA clones T3, T1 and OV.1

Regions of sequence identity are shaded. The sequencing primers used are indicated. Primer 1 is the reverse sequencing primer. Primers 2 and 6 are the universal sequencing primer. Primers 2, 3 and 5 were used for PCR analysis of the OV.1 cDNA (see text for details). ATG initiation codons, TAA termination codons and ATTAAA polyadenylation signals are underlined. In the OV.1 cDNA there was doubt as to whether the ATG at nt. 683 was the correct initiation codon (see text for details) and several alternative potential upstream initiation codons and the consensus translation initiation site at nt. 29 are also underlined.

T3 is a cDNA clone representing the full length *Drosophila* IMP dehydrogenase enzyme (see section 5.4).

T3	1.T3	1	39
T1	sequencing	AAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCC	
OV.1	primer →		
T3	40		89
T1	GCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGGCACGAGCACA		
OV.1			
T3	90		139
T1	CGCGCACTTGGACTCCAGCGGACTGCTCGAGAGGAATCTGTATCCACAAA		
OV.1			
T3	140		189
T1	AAGAAATCTGCAACAGGAGAAGATCCCAAAACAGGCGGTTTTTATCTCTA		
OV.1			
T3	190		239
T1	TACGGCTTGTTATTGTCAGCTCCCCGTTGTTTGTTGACTGTTTGTTTGGG		
OV.1			
T3	240		289
T1	TTGAACGCACGCGGAAA <u>ATGGAGAGCACGACAAAGGTGAAGGTGAACGGC</u>		
OV.1			
T3	290		339
T1	TTTGTGGAATCGACGTCGTCTTCAGCGGCGCCGGCAATCCAGACAAAGAG		
OV.1			
T3	340		389
T1	TACCACCGGATTCGATGCCGAGCTGCAGGTAGGGCTGAGTTGTAAGGAAC		
OV.1			
T3	390		439
T1	TGTTCCAGAACGGTGAGGGACTACCCTACAACGACTTTCTCATACTG <u>CCC</u>		
OV.1			
T3	440 T3 sequencing primer →		489
T1	<u>GGCTACATAGACTTCACCGCCGAGGAGGTCGATCTGAGTTCGCCACTGAC</u>		
OV.1			
T3	490		539
T1	CAAGTCGCTGACATTGCGAGCACCGCTGGTTAGTTCGCCCATGGACACGG		
OV.1			

	540		589
T3	TAACCGAATCGAGATGCCATGCCATGGCGCTGTGTGGTGGCATTGGCATC		
T1			
OV.1			
	590		639
T3	ATCCATCACAACTGCACGCCGGAATACCAGGCGTTGGAGGTGCACAAGGT		
T1			
OV.1			
	640		689
T3	TAAGAAGTACAAGCACGGCTTCATGCGCGACCCCTCGGTGATGTCGCCCA		
T1			
OV.1			
	690		739
T3	CGAATACGGTGGGGGATGTGTTGGAGGCGGGCGGAAGAACGGATTCACC		
T1			
OV.1			
	740		789
T3	GGCTATCCCGTCACCGAGAACGGCAAACCTGGCGGCAAGCTGCTGGGCAT		
T1			
OV.1			
	790		839
T3	GGTCACGTCGCGAGCAATTGACTTCCGCGAGAACCAGCCGGAGTCCCTGC		
T1			
OV.1			
	840		889
T3	TGGCCGACATCATGACCACCGAACTGGTCACCGCTCCCAATGGCATCAAT		
T1			
OV.1		2.OV.1 sequencing primer →	TTTGTTTACAGTAATTC
	890		939
T3	CTGCCCACGGAACACGCTATTCTCGAGAAGAGCAAGAAGGCAACTGCGAT		
T1			
OV.1	TTTTAATAAAAACATGTTAATCATTAGTTGGTTAAGCTTTCTGCTGCTGC		
			OV.1 translation start codon ?
	940		989
T3	TGTGAATCAGGCCGGCGAACTGGTGGCCATGATTGCCCCGCGGGATTTGA		
T1			
OV.1	TCCTGGCCGTAGTACTCGCTGTAATCATACTCCCTGAAGGGCACATCGTA		
			5. PCR primer →
	990		1039
T3	AGAAGGCCCGCTCCTACCCGAATGCCTCCAAGGACTCCAATACAAGGCTC		
T1			
OV.1	TGTTAGGTACCCGGAGTCGCGGGCTTGCTGAACGGCGACCATTAGACGCA		
	1040		1089
T3	CTCTGTCCGGCTGCCATCGGAACGCGATCGGAGGACAAGGGCTGTCGAGC		
T1			
OV.1	GGTGGCTCTTTTGACACAGGCCTGTCTTTGAGTAGCTAAGGACATCTCCG		

	1090	1139
T3	CCTGGCTCTGCTGGTGCACGCGTGGACGTTATCATCCTTGATTAT	
T1		
OV.1	CTGTGGGGCGAAATGAAGTCTCCAGGAGCTCCGTGTTGCGGTAGTCCAG	
	1140	1189
T3	CGCAGGGCAACTCCGTCATCCAGGTGGAGATGATCAAGTACATTAAGGAG	
T1		
OV.1	CACCAGGTATTCATCACGGCAAATGGGACAGGGATTTCCGGTGCTGATAC	
	1190	1239
T3	ACCTATCCCGAACTACAGGTCATTGGTGGCAATGTGGTAACACGTGCCCA	
T1		
OV.1	GGTTCTGGCGGATCGACTGCTTTCGTGTTTTAGCAGCGGCGCATACATGC	
	1240	1289
T3	GGCCAAAAATCTCATTGATGCCGGCGTCGATGGACTGCGTGTGGGCATGG	
T1	ATTGATGCCGGCGTCGATGGACTGCGTGTGGGCATGG	
OV.1	CCTTGTGGTTGCGCCGATACTGCGTCCACACAAAGTCTCGCCGTAGTTT	
	1290	1339
T3	GTTCCGGTTCCATCTGCATCACCCAGGAGGTGATGGCCTGCGGATGTCTT	
T1	GTTCCGGTTCCATCTGCATCACCCAGGAGGTGATGGCCTGCGGATGTCTT	
OV.1	GCTTGTAAGCGGCGCTCTTTAAGTATCGAATCGAGGTCTCCACTGGGATT	
	3. OV.1 sequencing primer →	
	1340	1389
T3	CAGGCCACTGCTGTCTACCAGGTGTCCACGTACGCTCGCCAATTCGGAGT	
T1	CAGGCCACTGCTGTCTACCAGGTGTCCACGTACGCTCGCCAATTCGGAGT	
OV.1	GGTTGCTGCATCCTTCGGGTCTGGCGCGGAACCACTAGCGTTATCCTCGC	
	1390	1439
T3	GCCAGTGATCGCGGACGGAGGCATCCATGCCATTGGACACATTGTGAAGG	
T1	GCCAGTGATCGCGGACGGAGGCATCCATGCCATTGGACACATTGTGAAGG	
OV.1	CTTCGCCTTCGGTGGCAGCCGCATTATTTTCTTCAGTCGAAGTCTCCGCC	
	1440	1489
T3	CTATTGCATTGGGAGCCAGCGCTGTAATGATGGGCTCCCTGCTGGCGGGC	
T1	CTATTGCATTGGGAGCCAGCGCTGTAATGATGGGCTCCCTGCTGGCGGGC	
OV.1	TTGCATCTCAAGGCGCTGGCCGTGTGAAGAGCTCCCCGGCTATTCCGGGT	
	← 4. T1 and T3 sequencing and PCR primer	
	1490	1539
T3	ACATCGGAGGCACCGGGCGAGTACTTCTTCTCCGACGGAGTGCGCCTTAA	
T1	ACATCGGAGGCACCGGGCGAGTACTTCTTCTCCGACGGAGTGCGCCTTAA	
OV.1	ACCTCGGAGGCACCGGGCGAGTACTTCTTCTCCGACGGAGTGCGCCTTAA	
	1540	1589
T3	GAAGTACCGCGGCATGGGCTCCCTGGAGGCCATGGAGCGCGGCGATGCCA	
T1	GAAGTACCGCGGCATGGGCTCCCTGGAGGCCATGGAGCGCGGCGATGCCA	
OV.1	GAAGTACCGCGGCATGGGCTCCCTGGAGGCCATGGAGCGCGGCGATGCCA	
	start of longest complete open reading frame	
	1590	1639
T3	AGGGCGCTGCCATGTCGCGCTACTACCACAACGAGATGGATAAAATGAAG	
T1	AGGGCGCTGCCATGTCGCGCTACTACCACAACGAGATGGATAAAATGAAG	
OV.1	AGGGCGCTGCCATGTCGCGCTACTACCACAACGAGATGGATAAAATGAAG	

	1640	1689
T3	GTGGCACAGGGCGTCAGTGGCAGCATCGTGGACAAGGGCAGTGTGCTGCG	
T1	GTGGCACAGGGCGTCAGTGGCAGCATCGTGGACAAGGGCAGTGTGCTGCG	
OV.1	GTGGCACAGGGCGTCAGTGGCAGCATCGTGGACAAGGGCAGTGTGCTGCG	
	1690	1739
T3	CTATCTACCCTATTTGGAGTGC GGACTGCAGCACAGTTGTCAGGACATCG	
T1	CTATCTACCCTATTTGGAGTGC GGACTGCAGCACAGTTGTCAGGACATCG	
OV.1	CTATCTACCCTATTTGGAGTGC GGACTGCAGCACAGTTGTCAGGACATCG	
	1740	1789
T3	GTGCTAACTCCATTAATAAAATTGAGGGATATGATCTATAATGGCCAACTG	
T1	GTGCTAACTCCATTAATAAAATTGAGGGATATGATCTATAATGGCCAACTG	
OV.1	GTGCTAACTCCATTAATAAAATTGAGGGATATGATCTATAATGGCCAACTG	
	1790	1839
T3	CGCTTCATGAAGCGCACCCATTCCGCTCAACTGGAAGGCAATGTCCATGG	
T1	CGCTTCATGAAGCGCACCCATTCCGCTCAACTGGAAGGCAATGTCCATGG	
OV.1	CGCTTCATGAAGCGCACCCATTCCGCTCAACTGGAAGGCAATGTCCATGG	
	T3 and T1 sequencing	
	1840	1889
	← primer	
T3	CCTCTTCAGCTACGAGAAGCGTCTCTTCTAAACATACGACGATGGCTGCGA	
T1	CCTCTTCAGCTACGAGAAGCGTCTCTTCTAAACATACGACGATGGCTGCGA	
OV.1	CCTCTTCAGCTACGAGAAGCGTCTCTTCTAAACATACGACGATGGCTGCGA	
	translation termination codon	
	1890	1939
T3	CGGAAACGGGCAGTGGAGGAGTGCCTGCCCCGTGCCAGCAGCAGCAGCAG	
T1	CGGAAACGGGCAGTGGAGGAGTGCCTGCCCCGTGCCAGCAGCAGCAGCAG	
OV.1	CGGAAACGGGCAGTGGAGGAGTGCCTGCCCCGTGCCAGCAGCAGCAGCAG	
	1940	1989
	T3 and OV.1 sequencing primer →	
T3	CAGCAGCAACACCATCAGCAGCAACAGCATCGCCACCGATCTGCAGCAGG	
T1	CAGCAGCAACACCATCAGCAGCAACAGCATCGCCACCGATCTGCAGCAGG	
OV.1	CAGCAGCAACACCATCAGCAGCAACAGCATCGCCACCGATCTGCAGCAGG	
	1990	2039
T3	CCCAGTCTCGACTACACCCGCCTATTTCGATCGCAAGTGAGAGAAGAAAA	
T1	CCCAGTCTCGACTACACCCGCCTATTTCGATCGCAAGTGAGAGAAGAAAA	
OV.1	CCCAGTCTCGACTACACCCGCCTATTTCGATCGCAAGTGAGAGAAGAAAA	
	2040	2089
T3	GAGAATCGAGCGCTCGAATAGCGCATCCTTTTCCCCTGGATTTTTTCCAA	
T1	GAGAATCGAGCGCTCGAATAGCGCATCCTTTTCCCCTGGATTTTTTCCAA	
OV.1	GAGAATCGAGCGCTCGAATAGCGCATCCTTTTCCCCTGGATTTTTTCCAA	
	2090	2139
T3	GGAATTCTGTAGGTTAAGTTTAATCGGAATCGCTGACGTTAAAGATTAC	
T1	GGAATTCT ← 6.T1 sequencing primer	
OV.1	GGAATTCTGTAGGTTAAGTTTAATCGGAATCGCTGACGTTAAAGATTAC	
	2140	2189
T3	GATTCCAATGTAGATTCCCCTTTTTGATGGCGGTGGAGGTTACCTCGAAT	
T1	GATTCCAATGTAGATTCCCCTTTTTGATGGCGGTGGAGGTTACCTCGAAT	
OV.1	GATTCCAATGTAGATTCCCCTTTTTGATGGCGGTGGAGGTTACCTCGAAT	

	2190	2239
T3	CGCCCGAATACATGTGTAATACTTAATCCTAATCTGTTCTGGTAGACAGA	
T1		
OV.1	CGCCCGAATACATGTGTAATACTTAATCCTAATCTGTTCTGGTAGACAGA	
	2240	2289
T3	ATCCTAATTATAAGTGTTTGCAGGTTACTTGATCTGTATAGAAAGTATAT	
T1		
OV.1	ATCCTAATTATAAGTGTTTGCAGGTTACTTGATCTGTATAGAAAGTATAT	
	2290	2339
T3	AATCTTCAACTATTCGCTATATATATCTAAAGCGCAACCGTTGGTTGGCC	
T1		
OV.1	AATCTTCAACTATTCGCTATATATATCTAAAGCGCAACCGTTGGTTGGCC	
	2340	2389
T3	GCTTACTTATTTTGCACGCTATATGTATATAGATGTTAATCTCTATTTGT	
T1		
OV.1	GCTTACTTATTTTGCACGCTATATGTATATAGATGTTAATCTCTATTTGT	
	2390	2439
T3	ATAGTGTTGGATGTTTACTATTATACCGAAGCTAGCTTCAAGTTTTAAAA	
T1		
OV.1	ATAGTGTTGGATGTTTACTATTATACCGAAGCTAGCTTCAAGTTTTAAAA	
	2440	2489
T3	CTCTGCCAAACTTATTAAATGTGAACTGAACTAACTGAAAAGTAAAAAA	
T1		
OV.1	CTCTGCCAAACTTATTAAATGTGAACTGAACTAACTGAAAAGTAAAAAA	
	polyadenylation signal	
	2490	2515
T3	AAAAAAAAAAAAAAAAAAAA	
T1		
OV.1	AAAAAAAAAAAAAAAAAAAA	

Figure 5.2

Nucleotide sequence and predicted amino acid sequence of cDNA T1

1 ATTGATGCCGGCGTCGATGGACTGCGTGTGGGCATGGGTTCCGGTTCCATCTGCATCACC
I D A G V D G L R V G M G S G S I C I T
61 CAGGAGGTGATGGCCTGCGGATGTCCTCAGGCCACTGCTGTCTACCAGGTGTCCACGTAC
Q E V M A C G C P Q A T A V Y Q V S T Y
121 GCTCGCCAATTCGGAGTGCCAGTGATCGCGGACGGAGGCATCCATGCCATTGGACACATT
A R Q F G V P V I A D G G I H A I G H I
181 GTGAAGGCTATTGCATTGGGAGCCAGCGCTGTAATGATGGGCTCCCTGCTGGCGGGCACA
V K A I A L G A S A V M M G S L L A G T
241 TCGGAGGCACCGGGCGAGTACTTCTTCTCCGACGGAGTGCGCCTTAAGAAGTACCGCGGC
S E A P G E Y F F S D G V R L K K Y R G
301 ATGGGCTCCCTGGAGGCCATGGAGCGCGGCGATGCCAAGGGCGCTGCCATGTGCGGCTAC
M G S L E A M E R G D A K G A A M S R Y
361 TACCACAACGAGATGGATAAAATGAAGGTGGCACAGGGCGTCAGTGGCAGCATCGTGGAC
Y H N E M D K M K V A Q G V S G S I V D
421 AAGGGCAGTGTGCTGCGCTATCTACCCTATTTGGAGTGCGGACTGCAGCACAGTTGTCAG
K G S V L R Y L P Y L E C G L Q H S C Q
481 GACATCGGTGCTAACTCCATTAATAAATTGAGGGATATGATCTATAATGGCCAACTGCGC
D I G A N S I N K L R D M I Y N G Q L R
541 TTCATGAAGCGCACCCATTCCGCTCAACTGGAAGGCAATGTCCATGGCCTCTTCAGCTAC
F M K R T H S A Q L E G N V H G L F S Y
601 GAGAAGCGTCTCTTCTAACATACGACGATGGCTGCGACGGAAACGGGCAGTGGAGGAGTG
E K R L F *
661 CCTGCCCCCTGTCCAGCAGCAGCAGCAGCAGCAACACCATCAGCAGCAACAGCATCGC
721 CACCGATCTGCAGCAGGCCCCACTGCTCGACTACACCCGCCTATTTCGATCGCAAGTGAGAG
781 AAGAAAAGAGAATCGAGCGCTCGAATAGCGCATCCTTTTCCCCTGGATTTTTTCCAAGGA
841 ATTC

Figure 5.3

Nucleotide sequence and predicted amino acid sequence of cDNA OV.1

The amino acid sequence indicated represents the longest continuous open reading frame (ORF). The underlined amino acid sequence represents the longest complete ORF with an ATG initiation codon. The true OV.1 translated region may extend beyond this ATG and encode a larger protein including the additional N-terminal sequence of 85 amino acids shown. The shaded 63 amino acid sequence is unique to the predicted OV.1 product. A likely alternative potential initiation codon, preceded by a consensus "AAAC" translation start site sequence, is underlined at nt. 32 (see text for details).

1 TTTGTTTACAGTAATTTCTTTTAATAAAAACATGTTAATCATTAGTTGGTTAAGCTTTC
61 TGCTGCTGCTCCTGGCCGTAGTACTCGCTGTAATCATACTCCCTGAAGGGCACATCGTAT
121 GTTAGGTACCCGGAGTCGCGGGCTTGCTGAACGGCGACCATTAGACGCAGGTGGCTCTTT
181 TGACACAGGCCTGTCTTTGAGTAGCTAAGGACATCTCCGCTGTGGGGCGAAATGAACTGC
241 TCCAGGAGCTCCGTGTTGCGGTAGTCCAGCACCAGGTATTCATCACGGCAAATGGGACAG
301 GGATTTCCGGTGCTGATACGGTTCTGGCGGATCGACTGCTTTCGTGTTTTAGCAGCGGCG
361 CATACATGCCCTTGTTGGTTGCGCCGATACTGCGTCCACACAAAGTCCTCGCCGTAGTTTG
421 CTTGTAAGCGGCGCTCTTTAAGTATCGAATCGAGGTCTCCACTGGGATTGGTTGCTGCAT
A A L F K Y R I E V S T G I G C C I
481 CCTTCGGGTCTGGCGCGGAACCACTAGCGTTATCCTCGCCTTCGCCTTCGGTGGCAGCCG
L R V W R G T T S V I L A F A F G G S R
541 CATTATTTTCTTCAGTCGAAGTCTCCGCTTGTCATCTCAAGGCGCTGGCCGTGTGAAGAG
I I F F S R S L R L A S Q G A G R V K S
601 CTCCCCGGCTATTCCGGGTACCTCGGAGGCACCGGGCGAGTACTTCTTCTCCGACGGAGT
S P A I P G T S E A P G E Y F F S D G V
661 GCGCCTTAAGAAGTACCGCGGCATGGGCTCCCTGGAGGCCATGGAGCGCGGCGATGCCAA
R L K K Y R G M G S L E A M E R G D A K
721 GGGCGCTGCCATGTCGCGCTACTACCACAACGAGATGGATAAAATGAAGGTGGCACAGGG
G A A M S R Y Y H N E M D K M K V A O G
781 CGTCAGTGGCAGCATCGTGGACAAGGGCAGTGTGCTGCGCTATCTACCCTATTTGGAGTG
V S G S I V D K G S V L R Y L P Y L E C
841 CGGACTGCAGCACAGTTGTCAGGACATCGGTGCTAACTCCATTAATAAATTGAGGGATAT
G L O H S C O D I G A N S I N K L R D M
901 GATCTATAATGGCCAACTGCGCTTCATGAAGCGCACCCATTCCGCTCAACTGGAAGGCAA
I Y N G O L R F M K R T H S A O L E G N
961 TGTCCATGGCCTCTTCAGCTACGAGAAGCGTCTCTTCTAACATACGACGATGGCTGCGAC
V H G L F S Y E K R L F
1021 GGAAACGGGCAGTGGAGGAGTGCCCTGCCCTGTCCAGCAGCAGCAGCAGCAGCAGCAACA
1081 CCATCAGCAGCAACAGCATCGCCACCGATCTGCAGCAGGCCCACTGCTCGACTACACCCG
1141 CCTATTCGATCGCAAGTGAGAGAAGAAAAGAGAATCGAGCGCTCGAATAGCGCATCCTTT
1201 TCCCCTGGATTTTTTCCAAGGAATTCTGTTAGGTTAAGTTTAATCGGAATCGCTGACGTT
1261 AAAGATTACGATTCCAATGTAGATTCCCCTTTTTTGATGGCGGTGGAGGTTACCTCGAATC
1321 GCCCGAATACATGTGTAATACTTAATCCTAATCTGTTCTGGTAGACAGAATCCTAATTAT
1381 AAGTGTTTGCAGGTTACTTGATCTGTATAGAAAGTATATAATCTTCAACTATTTCGCTATA
1441 TATATCTAAAGCGCAACCGTTGGTTGGCCGCTTACTTATTTTGCACGCTATATGTATATA
1501 GATGTTAATCTCTATTTGTATAGTGTGGATGTTTACTATTATACCGAAGCTAGCTTCAA
1561 GTTTTAAAACTCTGCCAACTTATTAAATGTGAACTGAACTAACTGAAAAGTAAAAAAA
1621 AAAAAAAAAAAAAA

Figure 5.4

Schematic of the relationship between the T3, T1 and OV.1 nucleotide and predicted amino acid sequences

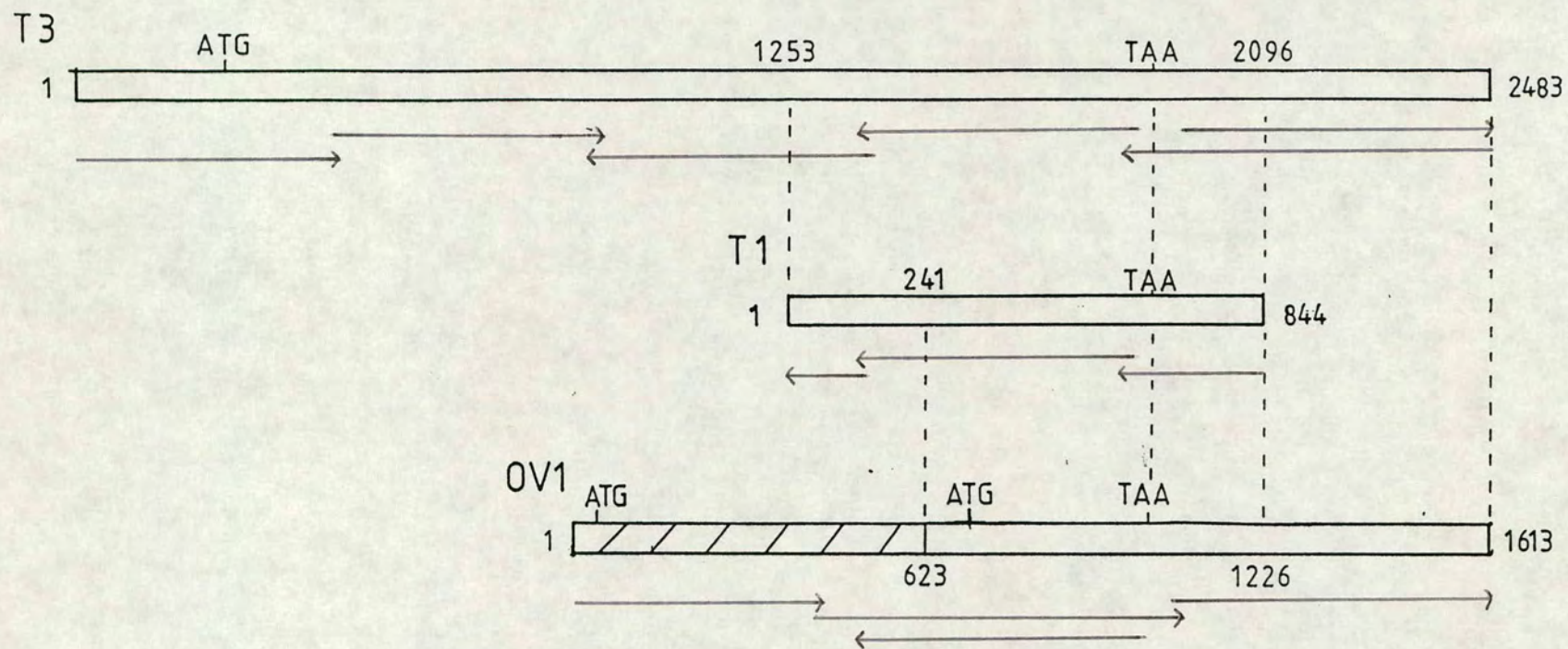
Open bar indicates regions of sequence identity.

Hatched bar indicates 5' sequence unique to cDNA OV.1.

ATG initiation codons and TAA termination codons are indicated.

The alternative potential upstream initiation codon of OV.1 at nt. 32 is indicated. Significant nucleotide positions are numbered.

The sequencing strategy is indicated by arrows.



extending from nt. 428 to nt. 997 and comprising the C-terminal 127 amino acids of the predicted T1 product plus a sequence of 63 N-terminal amino acids which are unrelated in sequence to the T1 product. Although this OV.1 ORF does not initiate with an ATG codon there are several potential start codons upstream of nt. 432 (Fig 5.1) and it seems possible that sequencing error has interrupted the reading frame. In particular, the furthest 5' of these other potential OV.1 start codons at nt. 32 (Fig 5.3) seems a likely candidate since it is preceded by the sequence "AAAC" which conforms fully with a derived *Drosophila* translation initiation site sequence (C\AAAA\C) (Cavener, 1987). Alternatively, the sequence as presented may be correct, in which case the longest complete ORF extends from the ATG at nt. 683 to nt. 997 and predicts a protein consisting of the C-terminal 105 amino acids of the predicted T1 product (Fig 5.3). The relationship between this protein and the predicted T1 product is shown in Fig 5.5. In order to resolve the question of the coding capacity of the OV.1 cDNA it will be necessary to obtain corroborative sequence from the complementary DNA strand in this region.

5.2 Is OV.1 an authentic cDNA ?

The sequence analysis of the OV.1 and T1 cDNA's which has been described indicated that they might represent two differentially spliced transcripts from the same gene, which share a region of 3' coding sequence. An alternative possibility is that the OV.1 cDNA does not actually represent a single mRNA sequence. In support of this hypothesis is the existence of poly (A) sequences and possible polyadenylation signals at both ends of this clone, suggesting that OV.1 may be a hybrid which arose from the co-ligation of two cDNA's during the cloning procedure. According to this theory, the point of overlap between the OV.1 and T1 sequences would represent the

Figure 5.5

Alignment of the predicted protein products of cDNA's T3, T1 and OV.1

The OV.1 product represented corresponds
to the longest complete open reading
frame present in the cDNA sequence.

	1		50
T3	MESTTKVKVN GFVESTSSA APAIQTSTT GFDAELQVGL SCKELFQNGE		
T1			
OV.1			
	51		100
T3	GLPYNDLIL PGYIDFTAEE VDLSSPLTKS LTLRAPLVSS PMDTVTESRC		
T1			
OV.1			
	101		150
T3	HAMALCGGIG IIHHNCTPEY QALEVHKVKK YKHGFMRDPS VMSPTNTVGD		
T1			
OV.1			
	151		200
T3	VLEARRKNGF TGYPVTENGK LGGKLLGMVT SRAIDFRENQ PESLLADIMT		
T1			
OV.1			
	201		250
T3	TELVTPNGI NLPTEHAILE KSKKATAIVN QAGELVAMIA RADLKKARSY		
T1			
OV.1			
	251		300
T3	PNASKDSNTR LLCPAAIGTR SEDKGCRALA LLVRNGVDVI ILDSSQGNSV		
T1			
OV.1			
	301		350
T3	IQVEMIKYIK ETYPELQVIG GNVVTRAQAK NLIDAGVDGL RVGMGSGSIC		
T1		IDAGVDGL RVGMGSGSIC	
OV.1			
	351		400
T3	ITQEVMACGC PQATAVYQVS TYARQFGVPV IADGGIHAIG HIVKAIALGA		
T1	ITQEVMACGC PQATAVYQVS TYARQFGVPV IADGGIHAIG HIVKAIALGA		
OV.1			
	401		450
T3	SAVMMGSLLA GTSEAPGEYF FSDGVRLKKY RGMGSLEAME RGDAKGAAMS		
T1	SAVMMGSLLA GTSEAPGEYF FSDGVRLKKY RGMGSLEAME RGDAKGAAMS		
OV.1		MGSLEAME RGDAKGAAMS	
	451		500
T3	RYYHNEMDKM KVAQGVSGSI VDKGSVLRYL PYLECGLQHS QDIGANSIN		
T1	RYYHNEMDKM KVAQGVSGSI VDKGSVLRYL PYLECGLQHS QDIGANSIN		
OV.1	RYYHNEMDKM KVAQGVSGSI VDKGSVLRYL PYLECGLQHS QDIGANSIN		
	501	537	
T3	KLRDMIYNGQ LRFMKRTHSA QLEGNVHGLF SYEKRLF		
T1	KLRDMIYNGQ LRFMKRTHSA QLEGNVHGLF SYEKRLF		
OV.1	KLRDMIYNGQ LRFMKRTHSA QLEGNVHGLF SYEKRLF		

junction of the two cDNA's rather than an intron \ exon boundary. Significantly, this would provide an explanation for the failure to detect a complete long ORF in the OV.1 sequence and other apparent anomalies in the data. If OV.1 does represent two mRNA's transcribed from different loci then the original pair of genomic clones (A22 E4 and A22 LD3) and the later set of clones (A22 LD4-A22 LD8) could derive from widely separated genomic regions, explaining the failure to isolate the intervening sequence. Also, according to this model, one of the transcripts represented in OV.1 could be the 1.0 kb ovary enriched mRNA transcribed from sequences upstream of p[lac Z ry⁺] which was originally detected by A-5.1 (Fig 3.6A) and recognised by OV.1 (Fig 4.1A) but not T1 (Fig 4.1B). The other transcript could be any of the various species recognised by OV.1 and T1 but not A-5.1. This theory could be tested by Northern blot analysis with a probe derived from the unique 5' portion of the OV.1 cDNA, which would be predicted to detect only the ovary enriched 1.0 kb transcript. Such a probe could be generated by the polymerase chain reaction (PCR) or by isolating a restriction fragment from this region of the OV.1 cDNA.

A number of features of the data are inconsistent with such a model and support the authenticity of the OV.1 cDNA. Firstly, the genomic fragment A-5.1 detects a number of transcripts in Northern hybridisation, which correspond in size and expression profile to transcripts detected by T1 (Fig 3.6A and 4.1B). Secondly, both OV.1 and lac Z hybridise to the same unique site on polytene chromosomes (Fig 3.3E and 3.3C). Thirdly, results obtained by PCR apparently confirm that OV.1 represents the product of a single locus. Oligonucleotide primers complementary to sequences on opposite DNA strands, at either side of the putative point of junction in OV.1 (primers 3 and 4 in Fig 5.1) were combined with OrR genomic DNA in a PCR reaction. The rationale of this experiment was that if OV.1 represented a single transcript, the two primers would

hybridise at adjacent sites on the genomic template and a PCR product would be obtained. Alternatively, if OV.1 were a hybrid clone no product would be expected. An added benefit of this strategy would be that if a PCR product were generated it should link the two sets of A22 genomic clones. In the event, a product was obtained. However, it was only about 1.5 kb in length and according to the restriction map data shown in Fig 3.5, the minimum size of intron expected was 8.3 kb. This suggested that the primers chosen for the PCR reaction were too close together and that the product did not span this intron, although it must include another intron(s) of about 1.3 kb since these primers lie only 186 bp apart in the OV.1 cDNA sequence (Fig 5.1). A repeat of this experiment with more widely separated primers, 529 bp apart in the OV.1 sequence (primers 4 and 5 Fig 5.1), in an attempt to amplify the missing intronic region, failed to generate a product. A reasonable explanation of this negative result could be that the distance separating the two primers on the genomic template was too great to permit amplification. The conclusion from this study that OV.1 is an authentic cDNA, rests on the assumption that the PCR product was genuine. Consistent with this, was the finding that the product hybridised strongly to OV.1 in a Southern blot (Fig 5.8A) but repeated attempts to confirm the identity of the PCR product by direct sequencing failed and I was unable to clone the product into a sequencing vector.

On balance, the data favour the view that OV.1 is an authentic cDNA. Moreover, the laboratory which constructed the ovary cDNA library concerned, has investigated several similar examples of clones possessing two poly (A) tails and concluded that they do represent single transcripts (D. Joseph, pers. comm.). Definitive proof in the case of OV.1 may require the isolation of the missing genomic DNA from a genomic library or the isolation of a cDNA of similar structure to OV.1 from a different cDNA library. Alternatively, primer extension (Calzone *et.al.*, 1987)

could be employed using a primer from the region common to OV.1 and T1 and an ovary poly (A)⁺ RNA template. Cloning and sequencing of the primer extension product(s) would establish whether OV.1 represents a single ovary transcript. A preliminary step which could provide further support for the authenticity of OV.1 would be to obtain genomic sequence from the proposed intron \ exon boundaries at the point of overlap between the OV.1 and T1 sequences in order to confirm the existence of splice donor and acceptor sequences at the predicted positions. This could be achieved using synthetic primers based on the OV.1 and T1 cDNA sequences and templates consisting of sub-cloned restriction fragments of the existing genomic clones i.e. fragment A-5.1 and one of the genomic sub-fragments recognised by T1 in Table 3.3.

(cDNA T1 also hybridises to position 8C-9A on polytene chromosomes -not shown).

5.3 The predicted products of T1 and OV.1 show high similarity to mouse and human inosine 5' monophosphate dehydrogenases (IMPD)

A search of the NBRF protein data base with the predicted protein sequences of T1 and OV.1 revealed a high similarity to the C-terminal portion of IMPD, the enzyme which catalyses the rate limiting step in *de novo* purine nucleotide synthesis (Weber *et.al.*, 1980; Weber, 1983), from various organisms (Fig 5.6). The closest similarity is to the mouse enzyme (72.6 % identical residues over the entire 205 amino acids of T1) followed by the human enzyme (72.1 % identical residues). The predicted 63 amino acid sequence unique to the OV.1 cDNA (Fig 5.3) showed no significant similarity to the IMPD sequences or any other sequence in the database and there was also no similarity between the untranslated regions of any of the IMPD sequences represented in Fig 5.6 and the corresponding regions of the T1 or OV.1 cDNA's.

Figure 5.6

Alignment of the the predicted product of cDNA T3 (*Drosophila*) with the amino acid sequences of IMPD from human, mouse, *Leishmania donovani* (Tryp.) and *E.coli*. The first amino acids of the predicted products of cDNA T1 (position 339) and cDNA OV.1 (position 439) are underlined. Regions of identity between the *Drosophila* product and the sequences of the other species are shaded.

	1				50
HumanMA	DYLISGGTGY	VP.....EDGL	TAQQLFASA.
MouseMA	DYLISGGTSY	VP.....DDGL	TAQQLFNCG.
Dros.	MESTTKVKVN	GFVESTSSSA	APAIQTKSTT	GFDAELQVGL	SCKELFQNG.
Tryp.MATNNAN	YRIKTIKDGC	TAEELFR.G.
E.coliMQSVTLICIM	PRQYLTTLV	EILPMLRIAK
	51				100
Human	DDLTYNDFLI	LPGFIDFIAD	EVDLTSALTR	KITLKTPLIS	SPMDTVTEAD
Mouse	DGLTYNDFLI	LPGYIDFTAD	QVDLTSALTK	KITLKTPLVS	SPMDTVTEAG
Dros.	EGLPYNDFLI	LPGYIDFTAE	EVDLSSPLTK	SLTLRAPLVS	SPMDTVTE.S
Tryp.	DGLTYNDFII	LPGFIDFGAA	DVNISGQFTK	RIRLHIPIVS	SPMDTITENE
E.coli	EALTFDDVLL	VPAHSTVLPN	TADLSTQLTK	TIRLNIPMLS	AAMDTVTEAR
	101				150
Human	MAIAMALMGG	IGFIHHNCTP	EFQANEVRKV	KNFEQGFITD	PVVLSPSHTV
Mouse	MAIAMALTGG	IGFIHHNCTP	EFQANEVRKV	KKYEQGFITD	PVVLSPKDRV
Dros.	RCHAMALCGG	IGIIHHNCTP	EYQALEVHKV	KKYKHGFMRD	PSVMSPTNTV
Tryp.	MAKTMALMGG	VGVLHNNCTV	ERQVEMVKS	KAYRNGFISK	PKSVPPNTPI
E.coli	LAIALAQEGG	IGFIHKNSI	EROAEEVRRV	KKHESGVVTD	PQTVLPPTTL
	151				200
Human	GDVLEAKMRH	GFSGIPITET	GTMGSKLVGI	VTSRDIDFLA	EKDHTTLLSE
Mouse	RDVFEAKARH	GFCGIPITDT	GRMGSRVLGI	ISSRDIDFLK	EEHDRFLEE
Dros.	GDVLEARRKN	GFTGYPVTEN	GKLGKLLGM	VTSRAIDF.R	ENQPESILAD
Tryp.	SNIIRIKEEK	GISGILVTEN	GDPHGKLLGI	VCTKDIDYVK	NKDTP..VSA
E.coli	REVKELTERN	GFAGYPVTE	E...NELVGI	ITGRDVRFVT	DLNQP..VSV
	201				250
Human	VMTPRIELVV	APAGVTLKEA	NEILQRSKKG	KLPIVNDCE	LVAIITARTDL
Mouse	IMTKREDLVV	APAGVTLKEA	NEILQRSKKG	KLPIVNEDE	LVAIITARTDL
Dros.	IMTT..ELVT	APNGINLPT	HAILEKSKKA	T.AIVNQAGE	LVAMIA RADL
Tryp.	VMTRREKMTV	ERAPIQLEEA	MDVLNRSRYG	YLPIVNEDE	VVNLCSSRDA
E.coli	YMTPKERLVT	VREGEAREVV	LAKMHEKRVE	KALVVDDEFH	LIGMITVKDF
	251				300
Human	KKNRDYPLAS	KDSQKQLLCG	AAVGTREDDK	.YR.LDLLTQ	AGVDVIVFHS
Mouse	KKNRDYPLAS	KDAKKQLLCG	AAIGTHEDDK	.YR.LDLLAL	AGVDVVVLDS
Dros.	KKARSYPNAS	KDSNTRLLCP	AAIGTRSEDK	GCRALALLVR	NGVDVILDS
Tryp.	VRARDYPHST	LDKSGRLICA	AATSTRPEDK	..RRVAALAD	VGVDVLVLDS
E.coli	QKAEAKPNAC	KDEQGRLRVG	AAVGAGAGNE	E..RVDALVA	AGVDVLLIDS
	301			*	350
Human	SQGNSVYQIA	MVHYIKQKYP	HLQVIGGNV	TAAQAKNLID	AGVDGLRVGM
Mouse	SQGNSIFQIN	MIKYIKEKYP	SLQVIGGNV	TAAQAKNLID	AGVDALRVGM
Dros.	SQGNSVIQVE	MIKYIKETYP	ELQVIGGNV	TRAQAKNLID	AGVDGLRVGM
Tryp.	SQGNTIYQIA	FIKWKSTYP	HLEVAVGNV	TQDQAKNLID	AGADGIRIGM
E.coli	SHGHSEGVLO	RIRETRAKYP	DLQIIGGNVA	TAAGARALAE	AGCSAVKVGI
	351				400
Human	GCGSICITQE	VMACGRPOGT	AVYKVAEYAR	RFGVPPIADG	GIQTVGHVVK
Mouse	GSIGSICITQE	VLACGRPOAT	AVYKVSEYAR	RFGVPVIADG	GIQNVGHIK
Dros.	GSIGSICITQE	VMACGCPOAT	AVYQVSTYAR	QFGVPVIADG	GIHAIGHIVK
Tryp.	GSIGSICITQE	VLACGRPOGT	AVYKVAQYCA	SRGVPCTADG	GLRQVGDICK
E.coli	GPGSICTTRI	VTGVGVPOIT	AVADAVEALE	GTGIPVIADG	GIRFSGDIAK

	401			*	450
Human	ALALGASTVM	MGSLLAATTE	APGEYFFSDG	VRLKKYRGMG	SLDPMEKS..
Mouse	ALALGASTVM	MGSLLAATTE	APGEYFFSDG	IRLKKYRGMG	SLDAMDKH..
Dros.	AIALGASAVM	MGSLLAGTSE	APGEYFFSDG	VRLKKYRGMG	SLEAMERGDA
Tryp.	ALAIGANCAM	LGGMLSGTTE	TPGEYFFKGG	VRLKVYRGMG	SLEAMSQ GK.
E.coli	AIAAGASAVM	VGSMLAGTEE	SPGEIELYQG	RSYKSYRGMG	SLGAMSKG..

	451				500
Human	.SSSQKRYF.	SEGDKVK.IA	QGVSGSIQDK	GSIQKFVPYL	IAGIQHGCOD
Mouse	.LSSQNRYF.	SEADKIK.VA	QGVSGAVQDK	GSIHKFVPYL	IAGIQHSCOD
Dros.	KGAAMSRYH	NEMDKMK.VA	QGVSGSIVDK	GSVLRYPYL	ECGLQHSCOD
Tryp.	..ESGKRYL.	SENEAVQ.VA	QGVSGNVVDK	GSAAKLIAYV	SKGLQQSAQD
E.coli	...SSDRYFQ	SDNAADKLVP	EGIEGRVAYK	GRLKEIIHQQ	MGGLRSCMGL

	501				550
Human	IGARSLSVLR	SMMYSGELKF	EKRTMSPQIE	GGVHGL....	HSYEKRLY..
Mouse	IGAKSLTQVR	AMTYSGELKF	EKRTSSAQVE	GGVHSL....	HSYEKRLF..
Dros.	IGANSINKLR	DMIYNGQLRF	MKRTHSAQLE	GNVHGL....	FSYEKRLF..
Tryp.	IGEISFDAIR	EKMYAGQVLF	SRRSPTAQGE	GGVHSL....	HSYEKKLFAA
E.coli	TGCGTIDELR	TKA.....EF	VRISGAGIQE	SHVHDVTITK	ESPNYRLGS.

	551
Human	..
Mouse	..
Dros.	..
Tryp.	KM
E.coli	..

5.4 Isolation of a cDNA encoding the full length *Drosophila* IMPD

Sequence analysis had indicated that T1 was a 5' truncated cDNA clone which was missing N-terminal coding sequence and was also lacking some 3' non-coding sequence (section 5.1.1). For these reasons, it remained doubtful whether this cDNA actually represented the 0.8 kb testis-specific transcript detected in the Northern analysis (Fig 4.1A and B). When a Stratagene *Drosophila* testis cDNA library in the vector lambda ZAP II (section 2.3.5.2) became available, this library was probed with the T1 cDNA with the aim of isolating a full length cDNA representing this transcript. In a screen of approximately 1.5×10^5 plaques only two positive clones were identified (T3 and T4). Both contained apparently identical inserts of about 2.5 kb. The size of the inserts and their homology to the T1 sequence, suggested that they might encode the full length *Drosophila* IMPD and this was confirmed by sequence analysis of the T3 clone (Fig 5.1). The T3 cDNA is 2483 nt. long excluding a poly (A) tail of 50 nt. Nucleotides 1493 to 2483 of T3 are identical to the last 990 nt. of OV.1 and nt. 1253 to 2096 of T3 constitute the entire sequence of T1 (Fig 5.1 and 5.4). A long open reading frame extends from the ATG at nt. 257 to nt. 1868, which encodes a protein of 537 amino acids (Fig 5.7). The relationship between the proteins encoded by the three cDNA's is illustrated in Fig 5.5).

The size of the predicted T3 protein product and its molecular weight (58 kD) is in good agreement with human and mouse IMPD (Natsumeda et.al., 1990; Tiedeman and Smith, 1991) and shows a high degree of sequence similarity to these two enzymes. Over its entire length it is 65.2 % identical to mouse IMPD and 64.8 % identical to the human enzyme. Over the C-terminal half of the enzyme the degree of identity is even higher (75.2 % between the *Drosophila* and mouse enzymes). There is also a high degree of sequence

Figure 5.7

Nucleotide sequence and predicted amino acid sequence of cDNA T3.

1 AAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCC
61 CCGGGCTGCAGGAATTCGGCACGAGCACACGCGCACTTGGACTCCAGCGGACTGCTCGAG
121 AGGAATCTGTATCCACAAAAAGAAATCTGCAACAGGAGAAGATCCCCAAAACAGGCGGTTT
181 TTATCTCTATACGGCTTGTTATTGTCAGCTCCCCGTTGTTTGTGACTGTTTGTGTTGGGT
241 TGAACGCACGCGGAAAATGGAGAGCACGACAAAGGTGAAGGTGAACGGCTTTGTGGAATC
M E S T T K V K V N G F V E S
301 GACGTCGTCTTCAGCGGCGCCGGCAATCCAGACAAAGAGTACCACCGATTTCGATGCCGA
T S S S A A P A I Q T K S T T G F D A E
361 GCTGCAGGTAGGGCTGAGTTGTAAGGAAGTGTTCAGAACGGTGAGGGACTACCCTACAA
L Q V G L S C K E L F Q N G E G L P Y N
421 CGACTTTCTCATACTGCCCCGCTACATAGACTTCACCGCCGAGGAGGTGCATCTGAGTTC
D F L I L P G Y I D F T A E E V D L S S
481 GCCACTGACCAAGTCGCTGACATTGCGAGCACCGCTGGTTAGTTCGCCCATGGACACGGT
P L T K S L T L R A P L V S S P M D T V
541 AACCGAATCGAGATGCCATGCCATGGCGCTGTGTGGTGGCATTGGCATCATCCATCACAA
T E S R C H A M A L C G G I G I I H H N
601 CTGCACGCCGGAATACCAGGCGTTGGAGGTGCACAAGGTTAAGAAGTACAAGCACGGCTT
C T P E Y Q A L E V H K V K K Y K H G F
661 CATGCGCGACCCCTCGGTGATGTCGCCACGAATACGGTGGGGGATGTGTTGGAGGCGCG
M R D P S V M S P T N T V G D V L E A R
721 GCGGAAGAACGGATTACACCGCTATCCCGTCACCGAGAACGGCAAACCTTGGCGGCAAGCT
R K N G F T G Y P V T E N G K L G G K L
781 GCTGGGCATGGTCACGTCGCGAGCAATTGACTTCCGCGAGAACCAGCCGGAGTCCCTGCT
L G M V T S R A I D F R E N Q P E S L L
841 GGCCGACATCATGACCACCGAACTGGTCACCGCTCCCAATGGCATCAATCTGCCCACGGA
A D I M T T E L V T A P N G I N L P T E
901 ACACGCTATTCTCGAGAAGAGCAAGAAGGCAACTGCGATTGTGAATCAGGCCGGCGAACT
H A I L E K S K K A T A I V N Q A G E L
961 GGTGGCCATGATTGCCCGCGCGGATTTGAAGAAGGCCCGCTCCTACCCGAATGCCTCCAA
V A M I A R A D L K K A R S Y P N A S K
1021 GGACTCCAATACAAGGCTCCTCTGTCCGGCTGCCATCGGAACGCGATCGGAGGACAAGGG
D S N T R L L C P A A I G T R S E D K G
1081 CTGTGAGCCCTGGCTCTGCTGGTGCACAACGGCGTGGACGTTATCATCCTTGATTCATC
C R A L A L L V R N G V D V I I L D S S
1141 GCAGGGCAACTCCGTCATCCAGGTGGAGATGATCAAGTACATTAAGGAGACCTATCCCGA
Q G N S V I Q V E M I K Y I K E T Y P E
1201 ACTACAGGTCAATTGGTGGCAATGTGGTAACACGTGCCCAGGCCAAAAATCTCATTGATGC
L Q V I G G N V V T R A Q A K N L I D A
1261 CGGCGTCGATGGACTGCGTGTGGGCATGGGTTCCGGTTCATCTGCATCACCAGGAGGT
G V D G L R V G M G S G S I C I T Q E V
1321 GATGGCCTGCGGATGTCCTCAGGCCACTGCTGTCTACCAGGTGTCCACGTACGCTCGCCA
M A C G C P Q A T A V Y Q V S T Y A R Q
1381 ATTCGGAGTGCCAGTGATCGCGGACGGAGGCATCCATGCCATTGGACACATTGTGAAGGC
F G V P V I A D G G I H A I G H I V K A
1441 TATTGCATTGGGAGCCAGCGCTGTAATGATGGGCTCCCTGCTGGCGGGCACATCGGAGGC
I A L G A S A V M M G S L L A G T S E A
1501 ACCGGGCGAGTACTTCTTCTCCGACGGAGTGCGCCTTAAGAAGTACCGCGGCATGGGCTC
P G E Y F F S D G V R L K K Y R G M G S
1561 CCTGGAGGCCATGGAGCGCGCGATGCCAAGGGCGCTGCCATGTCGCGCTACTACCACAA
L E A M E R G D A K G A A M S R Y Y H N

1621 CGAGATGGATATAAATGAAGGTGGCACAGGGCGTCAGTGGCAGCATCGTGGACAAGGGCAG
E M D K M K V A Q G V S G S I V D K G S
1681 TGTGCTGCGCTATCTACCCCTATTTGGAGTGC GGACTGCAGCACAGTTGTCAGGACATCGG
V L R Y L P Y L E C G L Q H S C Q D I G
1741 TGCTAACTCCATTAATAAAATTGAGGGATATGATCTATAATGGCCAACTGCGCTTCATGAA
A N S I N K L R D M I Y N G Q L R F M K
1801 GCGCACCCATTCCGCTCAACTGGAAGGCAATGTCCATGGCCTCTTCAGCTACGAGAAGCG
R T H S A Q L E G N V H G L F S Y E K R
1861 TCTCTTCTAACATACGACGATGGCTGCGACGGAACGGGCAGTGGAGGAGTGCCTGCCCC
L F *
1921 TGTCCAGCAGCAGCAGCAGCAGCAGCAACACCATCAGCAGCAACAGCATCGCCACCGATC
1981 TGCAGCAGGCCCACTGCTCGACTACACCCGCCTATTCGATCGCAAGTGAGAGAAGAAAAG
2041 AGAATCGAGCGCTCGAATAGCGCATCCTTTTTCCCCTGGATTTTTTCCAAGGAATTCTGTT
2101 AGGTTAAGTTTAAATCGGAATCGCTGACGTTAAAGATTACGATTCCAATGTAGATTCCCCT
2161 TTTTGATGGCGGTGGAGGTTACCTCGAATCGCCCGAATACATGTGTAATACTTAATCCTA
2221 ATCTGTTCTGGTAGACAGAATCCTAATTATAAGTGTTTGCAGGTTACTTGATCTGTATAG
2281 AAAGTATATAATCTTCAACTATTTCGCTATATATATCTAAAGCGCAACCGTTGGTTGGCCG
2341 CTTACTTATTTTTGCACGCTATATGTATATAGATGTTAATCTCTATTTGTATAGTGTGGA
2401 TGTTTACTATTATACCGAAGCTAGCTTCAAGTTTTTAAACTCTGCCAAACTTATTAAATG
2461 TGAAGTGAAGTAACTGAAAAGTAAAAAAAAAAAAA

similarity between the T3 product and IMPD from the trypanosome *Leishmania donovani* (49.3 % identity) and *E.coli* (38.8 % identity). An alignment of the predicted T3 product with the IMPD sequences of these four organisms is presented in Fig 5.6.

The high level of similarity between the sequences in Fig 3.14 constitutes compelling evidence that the T3 cDNA encodes the *Drosophila* IMPD enzyme, particularly since Southern analysis indicated that there are no other related sequences in the *Drosophila* genome (Fig 4.1C).

The results of the Northern analysis with the T1 cDNA (Fig 4.1B) suggest that the *Drosophila* IMPD gene gives rise to a number of differentially regulated transcripts. The OV.1 and T1 cDNA's may represent the 1.6 kb ovary enriched transcript and the 0.8 kb testis-specific transcript respectively, although for reasons discussed earlier this remains doubtful. Assuming these two cDNA's do genuinely represent transcripts from the *Drosophila* IMPD locus an obvious question concerns the function of the proteins they encode. The T1 cDNA apparently represents a transcript which encodes a protein consisting of a small number of unknown N-terminal amino acids plus the C-terminal 205 amino acids of IMPD (Fig 5.2 and 5.5). Until information becomes available regarding the location of functional domains within the IMPD protein, such as the binding sites of substrate and allosteric effectors, it is difficult to even speculate as to the possible function of the predicted T1 product. However, it may be significant that the T1 product includes the C-terminal portion of IMPD, which appears most highly conserved (Fig 5.6) and so presumably contains regions of functional significance. The OV.1 predicted protein product also comprises a C-terminal portion of IMPD of between 105 and 127 amino acids and may include an N-terminal sequence of amino acids which are unrelated to IMPD or to any other proteins in the database (Fig 5.3 and 5.5). Again, the sequence data offer little clue as to the possible function of such a protein.

5.5 Further investigation of the expression pattern of *Drosophila* IMPD

The sequencing results and the data from Northern hybridisation, suggest that the *Drosophila* IMPD gene gives rise to multiple transcripts with distinct expression patterns. Through a combination of further Northern analysis and *in situ* hybridisation techniques, a more comprehensive picture was obtained of the spatial and temporal regulation of this expression.

5.5.1 Northern analysis

Total RNA was extracted from a variety of *Drosophila* tissues and developmental stages and Northern blotted. The resulting filter was probed with the T1 cDNA, to derive a spatial and developmental profile of IMPD gene transcription. As Fig 5.8B shows, the level of the 2.5 kb transcript, which presumably encodes IMPD, is maximal in the adult ovary (lane 8) and it is also detected at an elevated level in embryos (lane 1). This result is consistent with the expression pattern of genes with a maternally encoded product which is delivered to the oocyte. However, since the RNA in lane 1 was extracted from embryos of mixed developmental stages (0-6 hours) it cannot be concluded with certainty that the transcript detected is of maternal origin and it will be necessary to repeat the experiment with RNA from staged embryos. It is also apparent from Fig 5.8B that a transcript of the same size is present at a reduced level in all other tissues and developmental stages examined.

The 0.8 kb transcript was detected only in adult testis RNA, where it appears more abundant than any other transcripts recognised by the T1 probe (Fig 5.8B lane 6). However, the existence of this transcript in the developing testes of earlier stages could not be ruled out, since it may be present at a level below the detectable threshold in larval and / or pupal total RNA. If such is the case, it may only be detectable in poly (A)⁺ RNA or RNA from

Figure 5.8

A: Southern blot of an Eco R1 digest of cDNA OV.1 cloned into plasmid pGEM-1 probed with the 1.5 kb PCR product.

The PCR product was generated with primers 3 and 4 (Fig 5.1). The strong band of hybridisation corresponds to the the cDNA insert at 1.6 kb.

B: Developmental profile of transcripts detected by cDNA T1

1 = 0-6 hour embryo

2 = third instar larva

3 = male pupa

4 = female pupa

5 = adult male carcass

6 = adult testes

7 = adult female carcass

8 = adult ovary

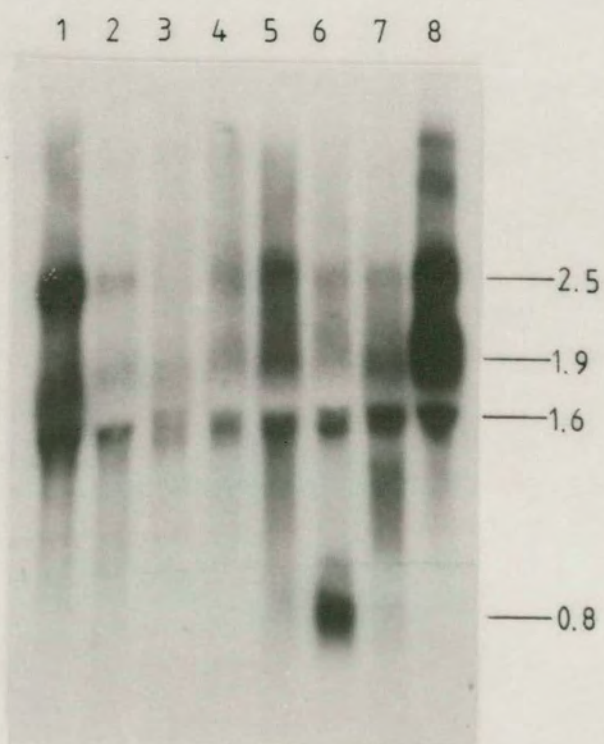
Approximately 10 μ g of total RNA were loaded per lane. The 1.9 kb transcript indicated is highly ovary enriched and is also present in embryo RNA at a level above that in the other tissues and stages examined.

The 2.5 kb transcript also appears enriched in ovary and egg RNA. The 0.8 kb transcript appears clearly testis-specific. Although some variability is apparent in the levels of the 1.6 kb, 1.9 kb and 2.5 kb transcripts in lanes 2-7 this is not seen consistently and may be attributable to unequal RNA loading or transfer.

A



B



larval and / or pupal total RNA. If such is the the case, it may only be detectable in poly (A)⁺ RNA or RNA from isolated testes.

Amongst the other transcripts recognised by the T1 probe, only the 1.9 kb species is clearly under tissue and developmental regulation. In common with the 2.5 kb transcript, it appears ubiquitous, but with an elevated abundance in ovaries, where it is the most abundant transcript and in early embryos. In contrast, the 1.6 kb transcript which may correspond to cDNA OV.1 is detected at a fairly constant level in all tissues and developmental stages examined. Two other minor transcripts of about 4 kb and larger are apparent in ovary and egg RNA, although these could represent unspliced precursors of smaller mRNA's.

5.5.2 Detection of IMPD transcripts *in situ*

The Northern hybridisation data had demonstrated that abundant transcripts from the IMP locus are present in the reproductive organs of both males and females, but was uninformative with respect to the particular tissue or cell types involved. In order to more precisely define the site of these transcripts, *in situ* hybridisation was performed to reproductive organs from both sexes. However, rather than utilising an autoradiographic procedure based on tritiated probes, I took advantage of a recently developed non-radioactive system, which permits signal detection in whole mounted tissues, as opposed to ultra-thin tissue sections (Tautz and Pfeifle, 1989). Probes are labelled with the steroid hapten, digoxigenin, and detected by an antibody / alkaline-phosphatase conjugate. Numerous successful applications of this technique have been reported, to material as diverse as *Drosophila* and *Xenopus* embryos (Tautz and Pfeifle, 1989; Hemmati-Brivanlou *et.al.* 1990).

5.5.2.1 Testis IMPD expression

Fig 5.9C is a photograph of adult male reproductive organs probed with the T1 cDNA. The hybridisation signal appears to be fairly uniformly distributed over the testes, accessory glands and other components, with no clear evidence of localisation to particular cell or tissue types except for a region of more intense staining in the apical region of one testis. Since IMPD does have a general "house-keeping" function in purine biosynthesis, such a transcript distribution is not unreasonable. However, the existence of the testis specific 0.8 kb transcript, which was detected on Northern blots, had implied that this gene may also have a specialised function in some aspect of male gametogenesis such as germ cell proliferation. If the 0.8 kb transcript does encode such a specific function, it seems logically probable that it be confined to a particular cell type within the reproductive system and in this regard it is surprising that no localised hybridisation signal was observed. One possible explanation for this result could be that any localisation of the 0.8 kb transcript was masked by the uniformly distributed mRNA species such as the full length IMPD transcript, which are also detected in testis RNA by the T1 probe (Fig 5.8B). However, in view of the relative abundance of the testis specific transcript, any localisation ought to be discernible against a uniform background signal. An alternative possibility is that the probe did not achieve equal penetration of all tissues, and indeed it has been reported that the testis sheath is probably impermeable (Dalby, B. pers.comm.). Consequently, transcripts expressed in the germline during spermatogenesis or in somatic cell types internal to the testis sheath may not be detected. Consistent with the hypothesis of unequal probe penetration is the increased signal level observed in a single testis in Fig 5.9C.

This limitation of the technique could be overcome by

Figure 5.9

***In situ* hybridisation to whole mounts of adult reproductive organs with digoxigenin labelled cDNA T1**

A: Ovary

X 10 magnification

nc = nurse cells

o = oocyte

A hybridisation signal is apparent from the earliest stages of oogenesis in the smallest egg chambers. The signal intensity increases with egg chamber maturation and is maximal in the nurse cells of stage 10B egg chambers, by which time staining is also apparent in the oocyte. Note the absence of staining from the indicated stage 11 oocyte.

B: Ovary subjected to post-hybridisation dehydration

X 10 magnification

This ovary was dehydrated with isopropanol and ethanol for long term storage. The treatment increases staining intensity with minimal loss of morphology.

The pattern of hybridisation mirrors that in A.

C: Male reproductive organs

X 10 magnification

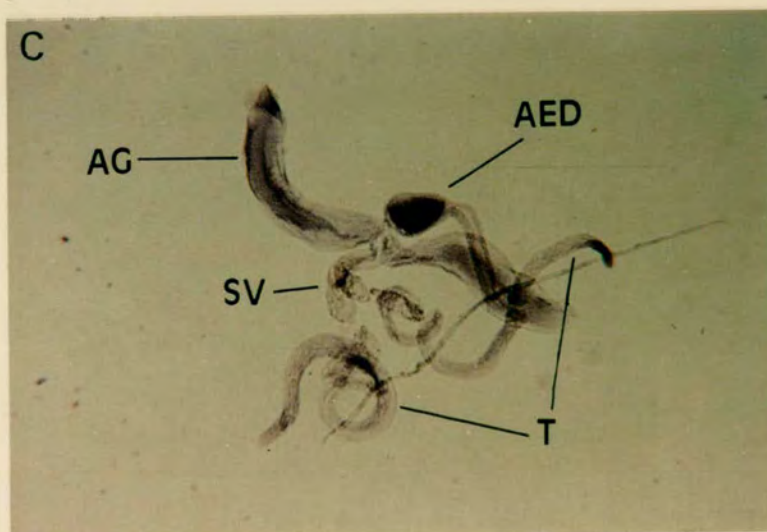
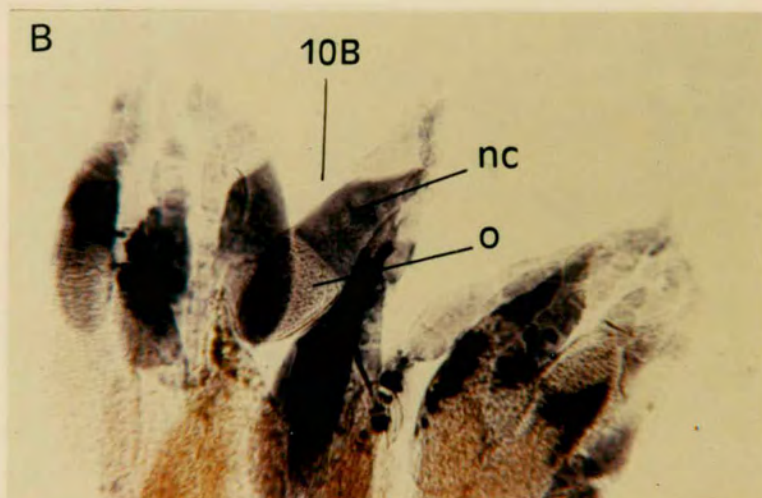
T = testis

AG = accessory gland

SV = seminal vesicle

AED = anterior ejaculatory duct

The staining appears fairly uniform throughout although there is an indication of increased intensity in the tip of the right-hand testis and the anterior ejaculatory duct.



the use of freeze fractured tissues to improve probe accessibility, the feasibility of which has been recently demonstrated (Dalby, B pers.comm.).

5.5.2.2 Ovarian IMPD expression

In adult ovaries, the T1 probe hybridises predominantly to germline derived cell types (Fig 5.9A and B). A relatively low level of transcripts is detected in the nurse cells of early egg chambers and this increases progressively with egg chamber maturation. The most intense hybridisation signal is observed in the nurse cells at stage 10B of oogenesis and by this stage transcripts are also clearly detected in the oocyte. This pattern of gene expression is consistent with the data from Northern analysis which revealed a high level of transcripts in ovaries and embryos (Fig 5.8B) and resembles the expression of other genes with established maternal functions such as *bicoid* (Berleth et.al., 1988) and *vasa* (Lasko and Ashburner, 1990).

A notable feature of the *in situ* results is the absence of a hybridisation signal from some later stage oocytes (Fig 5.9A). This seems at odds with the proposal that maternally supplied transcripts from the IMPD gene are stored for later use by the early embryo. However, as in the case of the testis experiment, this result may be a reflection of uneven probe penetration, in this case due to the chorionic layers, which are deposited in the later oogenic stages and cannot be easily removed without loss of tissue morphology. A similar failure to consistently detect transcripts by this method in late stage oocytes, has been observed by other members of this laboratory (Grimes, B. and Xiang, N pers.comm.) although there are also examples in the literature of transcripts being successfully detected in later stage egg chambers, using modifications to the original whole mount *in situ* protocol designed to improve probe accessibility (e.g. Kim-Ha et.al., 1991).

In Northern blots of ovary RNA T1 detects several transcripts of similar abundance (Fig 5.8B) and the pattern of hybridisation observed in the whole mount *in situ* experiment therefore reflects their combined distributions. It is consequently impossible to determine whether any of these transcripts are subject to independent temporal or spatial regulation. This problem can hopefully be overcome by identifying either genomic or cDNA probe sequences specific for the various transcripts on a Northern blot, which could then be used for *in situ* hybridisation experiments.

5.5.2.3 Embryo *in situ*

The same *in situ* hybridisation technique was applied with the aim of determining the embryonic distribution of IMPD transcripts. In embryos of all stages the immunodetection indicated an identical, uniform distribution of IMPD transcripts. However, the homogeneous colouration which was observed was indistinguishable from the pattern observed in negative control experiments using digoxigenin labelled plasmid (pGEM-1) as a probe so it seemed possible that the result obtained with the IMPD probe reflected non-specific background hybridisation rather than IMPD transcript distribution. Again, this may have been a consequence of poor probe penetration and it is intended to repeat the experiment with modifications aimed at ensuring the removal of the chorionic and vitelline layers.

CHAPTER SIX

FINAL DISCUSSION

6.1 Functions of IMPD

IMP dehydrogenase has a key role in the de novo synthesis of purine nucleotides. It catalyses the rate limiting step in the production of guanosine monophosphate (GMP), which consists of the oxidation of inosine monophosphate (IMP) to xanthine monophosphate (XMP) in the presence of the co-factors nicotinamide adenine dinucleotide (NAD) and adenosine triphosphate (ATP) (Fig 6.1) (Weber *et.al.*, 1980; Weber, 1983). The reaction is of particular metabolic significance since IMP is at the branchpoint of the pathways of adenosine monophosphate (AMP) and GMP synthesis and IMPD thus has a potential regulatory function in determining the relative cellular concentrations of these two nucleoside phosphates.

6.2 IMPD is implicated in tumorigenesis

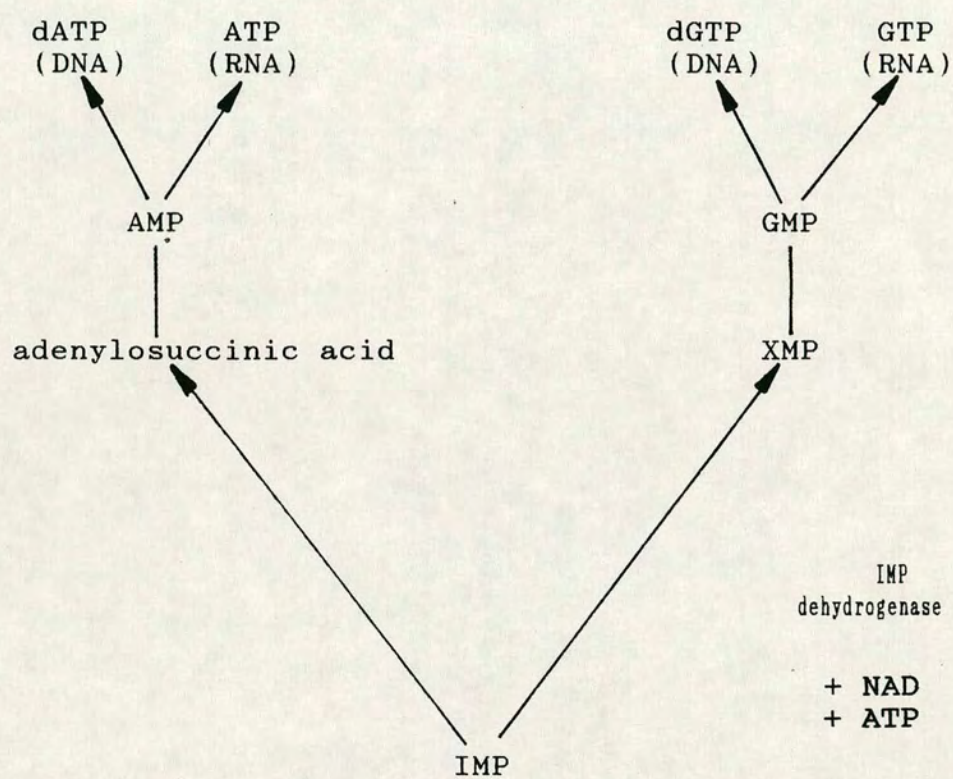
The pivotal role of this enzyme in providing the precursors of DNA and RNA biosynthesis is reflected in a correlation between IMPD activity and cellular proliferation. Inhibition of the enzyme by mycophenolic acid causes an abrupt cessation of DNA synthesis (Cohen *et.al.*, 1981; Duan and Sadee, 1987) and blockage of the cell cycle at the G₁-S interface (Cohen and Sadee, 1983). In healthy tissues an association has been observed between the level of IMPD activity and the rate of cell division (Jackson and Weber, 1975; Jackson *et.al.*, 1977; Cooney *et.al.*, 1983; Profitt *et.al.*, 1983; Shimura *et.al.*, 1983) and a number of studies have implicated the enzyme in the process of neoplastic transformation. A range of rat hepatomas was found to exhibit levels of IMPD activity which were correlated with growth rate and in the most rapidly growing tumours the level was up to eleven-fold greater than the value for normal liver. Furthermore, in comparison with the correlation between rate of cell division and IMPD activity seen in healthy tissue, the

Figure 6.1

The biochemical pathway of *de novo* purine nucleotide synthesis

IMP = inosine monophosphate

XMP = xanthine monophosphate



elevated activity in the tumours was disproportionately high, suggesting a possible causal link between IMPD activity and transformation or tumour progression (Weber, 1983; Jackson and Weber, 1975). Additional evidence of the significance of this enzyme in cancer has come from the observation that as rat hepatoma cells enter the proliferative programme there is an increased channeling of IMP into guanylate biosynthesis at the expense of adenylate (Natsumeda *et.al.*, 1988). A suggestion that the role of IMPD in transformation may be mediated by the activation of cellular proto-oncogenes came from an examination of the effects of the inhibitor of IMPD, tiazofurin, on a human leukaemia cell line. The induction of the erythroid differentiation programme in these cells by tiazofurin was found to be preceded by a decrease in IMPD activity and a down-regulation of the proto-oncogene c-Ki-ras (Olah *et.al.*, 1988).

Its evident importance in cancer has made IMPD a promising target in chemotherapy. The nucleoside tiazofurin, when introduced to cells, is converted to the NAD analogue, thiazole-4 carboxamide adenine dinucleotide (TAD), an analogue of NAD. TAD selectively blocks the activity of IMPD (Cooney *et.al.*, 1982; Jayaram *et.al.*, 1982) by binding to the NADH inhibition site (Yamada *et.al.*, 1988). A broad spectrum of tumours is suppressed by tiazofurin, including Lewis lung carcinoma, which is refractory to most drugs and it has proven effective against human leukaemia in clinical trials (Weber *et.al.*, 1988; Tricot *et.al.*, 1989).

6.3 *Drosophila* IMPD

This thesis describes the isolation of a cDNA clone encoding *Drosophila* IMPD. The molecular weight and amino acid sequence of the predicted protein product agree closely with data on this enzyme from a range of phylogenetically diverse invertebrate and vertebrate

organisms, indicating a high degree of conservation through evolution, which presumably reflects the ubiquity of the metabolic pathway which IMPD catalyses. Studies on the expression of the *Drosophila* gene revealed a complexity that would not have been predicted from the enzyme's known functions. A number of overlapping transcripts with different sizes and expression profiles are produced. At least three of these transcripts are present in both gonadal and carcass tissue of male and female flies, including the one most likely to encode full length IMPD, on account of its size (2.5 kb). However, there is considerable enrichment for these three transcripts in adult ovaries and *in situ* hybridisation to whole mounted female gonads with IMPD probe sequences permitted the localisation of the ovarian transcripts to the nurse cells and late stage oocytes. This distribution of transcripts suggested a maternal role for this gene in *Drosophila*, supplying a product to the developing zygote, and this conclusion was reinforced by the detection of high levels of transcripts in embryos.

The high levels of nurse cell IMPD mRNA are consistent with the heavy demand for nucleotides and deoxynucleotides which the polyploidisation and elevated transcriptional activity in this cell type must impose. Presumably, the maternally supplied mRNA is necessary to provide for DNA replication during the rapid nuclear proliferation which characterises early embryogenesis, prior to the onset of zygotic transcription. A similar pattern of expression has been reported for two other *Drosophila* genes with an established role in cell cycle regulation which encode cyclin A and cyclin B (Whitfield et.al.,1989). Both of these genes encode abundant maternal mRNA's which are supplied to the embryo. These transcripts are initially uniformly dispersed but as development proceeds there is a marked redistribution. During the 8th and 9th mitotic divisions, cyclin A transcripts migrate towards the cortex in association with the zygotic nuclei. At the same time,

cyclin B transcripts become even more densely concentrated around the somatic nuclei than the cyclin A transcripts as well as becoming highly concentrated at the embryo posterior, in the region of pole cell formation; a process probably requiring a component of the polar granules. (Raff et.al., 1990). The cyclin B transcripts are incorporated into the pole cells and continue to be detected in these cells as they move dorsally and anteriorly during germ band elongation (Whitfield et.al., 1989; Raff et.al., 1990). Following the disappearance of the maternal transcripts, the two cyclin genes display distinct patterns of zygotic expression during subsequent development, in tissues which remain mitotically active. In the case of cyclin A, transcripts are detected in the larval brain lobes and imaginal discs. The predominant zygotic cyclin B transcript is smaller than the maternally supplied message and was reported to be testis-specific (Whitfield et.al., 1989) but more recent data indicate its presence in the male and female germlines (Dalby and Glover, 1992).

Although the preliminary *in situ* hybridisation experiments reported here revealed no embryonic localisation of IMPD mRNA the results were inconclusive due to technical difficulties and could not be informative with respect to the distributions of individual transcripts. In view of the enzyme's association with cell division, it would be interesting to repeat the experiment with probes specific for the various IMPD transcripts to ascertain whether any of them were subject to localisation, as shown for the cyclin transcripts. Whitfield et.al (1989) proposed that the cyclin transcripts may contain localisation signals directing their embryonic redistribution and it has since been shown that the maternal cyclin B transcript possesses within its 393 nt. 3' untranslated region, a sequence which is necessary for its pole cell localisation and its association with the somatic nuclei (Dalby and Glover, 1992). The embryonic localisation of maternal *bcd* transcripts has similarly been shown to depend on signals

within its untranslated trailer region (MacDonald and Struhl, 1988). In this connection, it may be significant that the IMPD cDNA includes an extended 3' untranslated region, although it shows no significant sequence similarity to the corresponding region of the cyclin B message or the *bcd* message.

It has been shown in the present study that in addition to the transcripts discussed above, the *Drosophila* IMPD gene produces a testis specific transcript of 0.8 kb. The coding capacity of this transcript was not determined with certainty but it appears that it may encode a protein which consists of around 200 C-terminal amino acids of the full length enzyme. The possible functions of such a truncated product remain obscure. As information on the functional domains of the IMPD enzyme becomes available and the sequence of the 0.8 kb transcript is confirmed, it may be possible to deduce a function for the protein it encodes. An initial requirement will be to accurately determine the cell type specificity (germline or somatic) of the transcript by an improved in situ hybridisation method, should a probe specific for this transcript be obtainable. Alternatively, Northern analysis of flies which lack a germline could be used to distinguish germline or somatic specificity of the transcript. Flies of this phenotype can be obtained as progeny of mothers which are homozygous for an hypomorphic mutation of the posterior group gene *tud* (Boswell and Mahowald, 1985) or *osk*⁵⁰¹ (Lehmann and Nüsslein-Volhard, 1986).

Several features of germ cell proliferation distinguish it from the mitotic cell division of somatic tissues. In the imaginal discs for example, there are a finite number of cell divisions after which the morphogenesis of adult structures occurs through cell enlargement and migration. In contrast, stem cell proliferation continues throughout adult life. Furthermore, gametogenesis involves syncytial nuclear divisions and meiosis. It seems reasonable to propose therefore, that a

specific form of IMPD could be required during spermatogenesis on account of these differences, particularly if IMPD performs a regulatory role as suggested earlier. A similar argument was forwarded by Whitfield *et.al.* (1989) to explain the apparent specificity to the male germline of the zygotic cyclin B transcript. A corollary of this argument might be that oogenesis should also require such a gene product although there was no evidence from Northern hybridisation of an absolutely ovary-specific transcript from this gene. However, spermatogenesis does differ from oogenesis in several important respects and might reasonably have unique requirements. For example, in the male, a stem cell gives rise to 64 haploid gametes which are transcriptionally inactive, whereas in females it produces 15 transcriptionally active polyploid nurse cells and an oocyte.

6.4 Future developments

A principal future objective in the analysis of the *Drosophila* IMPD gene which has been partially characterised in the course of this work, must be to ascertain the coding capacities of all the various transcripts which it generates. This will involve the isolation and sequencing of cDNA clones corresponding to these transcripts. In parallel with this work it would be desirable to clone the remaining genomic sequences which constitute this locus. By comparing the structure and sequence of the cDNA and genomic clones it will be possible to determine the transcriptional and/or post-transcriptional mechanisms by which the multiple mRNA's are produced. S1 nuclease mapping (Berk and Sharp, 1978) of the cloned genomic region will assist in deriving the transcriptional organisation of the gene and having determined the sequence of the cDNA clones, the 5' ends of the various transcripts can be determined by primer extension analysis of poly (A)⁺ RNA (Calzone *et.al.*,

1987).

The availability of the genomic region will facilitate a search for cis-acting elements directing the observed transcription pattern. These might be expected to include an enhancer sequence mediating elevated transcription in the nurse cells and it may also be discovered that the expression of the 0.8 kb testis-specific transcript is under the control of a separate enhancer. The identification of these elements will initially involve P-element transformation experiments with constructs comprising sections of flanking DNA linked to a suitable reporter gene - a technique successfully applied to a number of *Drosophila* genes including the yolk protein genes and the $\beta 2$ tubulin gene (discussed in chapter one) and the *white* gene (reviewed in Hazelrigg, 1987). Should a testis-specific enhancer be found associated with the IMPD gene it will be interesting to look for sequence similarity with the $\beta 2$ UE1 element upstream of the $\beta 2$ tubulin gene. The possibility of translational control in the case of the testis-specific 0.8 kb IMPD transcript could also be investigated if an antibody were obtained against the protein product, as reported for the gene *mst 87F* (see chapter one) and the untranslated portion of the 0.8 kb transcript could be examined for any sequence similarity to the translational control element in *mst 87F*. Since there is presumably no demand for nucleotide synthesis following the cessation of mitosis and transcriptional activity during spermatogenesis, such translational regulation of IMPD expression might be unlikely. However, translational control could also be investigated in the ovary, where a proportion of the IMPD mRNA must remain untranslated for export to the developing oocyte and the availability of an antibody would also be useful in determining the sub-cellular localisation of the IMPD protein.

In considering the functions of the various transcripts produced by the *Drosophila* IMPD gene it would

be of value to more precisely determine all their respective tissue specificities. This could be accomplished by Northern analysis with isolated tissue types or by in situ hybridisation, with probes specific for each transcript or antibodies to the protein product(s). As more information becomes available, it will be possible to compare the modes of regulation employed in other organisms with the situation in *Drosophila*. In humans, it has already been shown that there exist two IMPD genes, one generating a 2.3 kb transcript and the other generating a 3 kb transcript (Natsumeda et.al., 1990). The transcripts differ in the length of their non coding sequences but both encode proteins of 514 amino acids which are 84% identical. Northern analysis revealed that there was a pronounced difference between normal leukocytes and ovarian tumours in the relative proportions of the two RNA species. Mammalian and bacterial IMPD's have a tetrameric structure (Yamada et.al., 1988; Gilbert et.al., 1979) and Natsumeda et.al., 1990) noted that the existence of the two IMPD genes in humans offered the potential for the production of tissue specific isoforms which might have different catalytic or allosteric properties. In support of this possibility, it has been shown by Okada et.al (1983) that IMPD isolated from rat liver and a rat sarcoma cell line differed in their response to the allosteric effectors XMP, GMP, AMP and NAD, a finding which also suggests that two IMPD genes may exist in the rat.

In view of this evidence for possible tissue specific isoenzymes in other species, one could speculate that the multiple IMPD transcripts in *Drosophila* represent an alternative to gene duplication for the production of such isoenzymes, involving the production of different enzyme sub-units from the same gene.

The housekeeping role of IMPD makes it likely that loss of function mutations abolishing the enzyme activity would be homozygous lethal. However, it may be possible to generate EMS mutations affecting specific transcripts from

the gene, permitting the loss of function phenotype of for example the testis specific product to be investigated. Existing deficiencies in the region 38C (Lindsley and Grell, 1968) will be useful in this undertaking or more localised deficiencies could be generated by imprecise excision of p[lac Z, ry⁺]. Also, recently developed technology, which might be usefully applied to the IMPD gene, offers the potential of obtaining cell type specific over-expression of any cloned gene in *Drosophila* (Wilson et.al., 1990; A. Brand and N. Perrimon, pers.comm. in Freeman, 1991). This involves placing the gene downstream of a yeast upstream activator sequence (UAS) and transforming into a line of flies which expresses the cognate yeast transcriptional activator GAL4, exclusively in the desired cell type. The system is based on the production of so called second generation enhancer-trap lines in which the β -galactosidase reporter gene is replaced by a β -galactosidase / GAL-4 fusion gene. Cell type specific expression of the fusion gene can be detected by staining for β -galactosidase activity. Transposant lines exhibiting the required expression pattern can then be crossed to flies carrying a second experimental transgene downstream of a yeast UAS and this cross will yield progeny which express the experimental gene in a pattern matching the observed β -galactosidase activity. It has already been shown that GAL-4 can activate transcription of a reporter gene downstream of a UAS in a variety of *Drosophila* tissues (Fischer et.al., 1988) and the feasibility of a similar system employing the Herpes Simplex Virus 1 transcriptional activator VP 16, has been demonstrated in mice (Byrne and Ruddle, 1989). Numerous applications could be envisaged for such a system, including the study of development by cell type specific ablation. This could be achieved by expression of cytotoxins in a particular cell type, a technique which has been successfully applied to the ablation of *Drosophila* photoreceptor cells (Kunes and Steller, 1991). The ability to obtain ectopic or

overexpression of developmentally interesting genes would also be of great value. In the case of the *Drosophila* IMPD gene, the application of this system might be of relevance to cancer research since it could potentially reveal whether excessive levels of the enzyme were sufficient to trigger neoplastic transformation by overexpressing IMPD in a restricted group of cells.

The route to the isolation of the IMPD gene, which is described in this thesis, began with an enhancer-trap line exhibiting testis specific β -galactosidase expression. Although the IMPD gene does produce a testis-specific transcript and may have a corresponding enhancer element, it was impossible to say whether or not this element was responsible for the observed staining pattern. Only when the location of the IMPD gene and its regulatory elements with respect to the P-element promoter in p[lac Z,ry⁺] has been determined, could this question be addressed and the data as they stand leave a strong possibility that the isolation of the gene was fortuitous.

REFERENCES

- Allen, R., O'Brien, D. and Eddy, E. (1988). *Mol.Cell.Biol.* 8: 828-832
- Ambrosio, L., Mahowald, A.P. and Perrimon, N. (1989). *Development* 106: 145-158
- Amrein, H., Maniatis, T. and Nöthiger, R. (1990). *EMBOJ* 9: 3619-3629
- Anderson, K.V., Bokla, L. and Nüsslein-Volhard, C. (1985). *Cell* 42: 791-798
- Anderson, K.V., Jürgens, G. and Nüsslein-Volhard, C. (1985). *Cell* 42: 779-789
- Anderson, K.V. (1987). *Trends In Genetics* 3: 91-97
- Anderson, K.V. and Nüsslein-Volhard, C. (1986). *Symp.Soc.Dev.Biol.* 44: 177-194
- Anderson, K.V. (1989). *Drosophila The Maternal Contribution*. In: *Genes and Embryos*. D.M. Glover and B.D. Hames (eds). IRL Press: 1-37
- Artavanis-Tsakanos, S. and Simpson, P. (1991). *TIG* 7: 403-408
- Ashburner, M. (1989). *Drosophila A Laboratory Handbook*. Cold Spring Harbor Laboratory Press.
- Baker, B.S. and Belote, J. (1983). *Ann. Rev. Gen.*, 17: 345-375
- Baker, B.S. and Ridge, K. (1980). *Genetics* 94: 383-423
- Bakken, A. (1973). *Dev.Biol.* 33: 100-122
- Barnett, T., Pachl, C., Gergen, J.P. and Wensink, P.C. (1980). *Cell* 21: 729-738
- Bell, A. (1954). *Genetics* 29: 958-959
- Bell, L., Horabin, J., Schedl, P. and Cline, T. (1991). *Cell* 65: 229-239
- Bell, L., Maine, E., Schedl, P. and Cline, T. (1988). *Cell* 55: 1037-1046
- Bellen, H.J., O'Kane, C.J., Wilson, C., Grossniklaus, U., Kurth Pearson, R. and Gehring, W.J. (1989). *Genes and Development* 3: 1288-1300
- Belote, J. and Baker, B. (1982). *Proc.Natl.Acad.Sci. USA* 79: 1568-1572
- Belote, J. and Baker, B. (1983). *Dev.Biol.* 95: 512-517

- Berk, A.J. and Sharp, P.A. (1978). Proc.Natl.Acad.Sci. USA 75: 1274-1278
- Berleth, T., Burri, M., Thoma, G., Bopp, D., Richstein, S., Frigerio, G., Noll, M., Nüsslein-Volhard, C. (1988). EMBOJ. 7: 1749-1756
- Bier, E., Vaessin, H., Shepherd, S., Lee, K., McCall, K., Barbel, S., Ackerman, L., Carretto, R., Uemura, T., Grell, E., Jan., L. and Jan, Y. (1989). Genes Dev. 3: 1273-1287
- Birnboim, H.C. and Doly, J. (1979). Nucleic Acids Research 7: 1513-1523
- Boggs, R., Gregor, P., Iddriss, S., Belote, J.M., and McKeown, M. (1987). Cell 50: 739-747
- Bonaccorsi, S., Pisano, C., Puoti, F. and Gatti, M. (1988). Genetics 120: 1015-1034
- Boswell, R.E. and Mahowald, A.P. (1985) Cell 43: 97-104
- Bownes, M. and Blair, M. (1986). J.Insect.Physiol. 32: 493-501
- Bownes, M. and Dale, L., (1982). In: " A Handbook of *Drosophila* Development" ed. R. Ransom, Elsevier Biomedical Press: 31-48
- Bownes, M. and Hames, B.D. (1978). Febs Lett. 96: 327-330
- Bownes, M. and Nöthiger, R. (1981). Mol. Gen. Genet. 182: 225-228
- Bownes, M., Steinmann-Zwicky, M. and Nöthiger, R. (1990). EMBOJ 9: 3975-3980
- Brennan, M., Wiener, A., Goralski, T. and Mahowald, A. (1982). Dev.Biol. 89: 225-236
- Bridges, C. (1921). Science 54: 252-254
- Bridges, C. (1925). Am.Nat. 59: 127-137
- Brousseau, G. (1960). Genetics 45: 257-274
- Bryant, P. and Hsei, B. (1977). Amer.Zool. 17: 595-611
- Burke, T., Waring, G., Popodi, E. and Minoo, P. (1987). Dev.Biol. 124: 441-450
- Burtis, K. and Baker, B. (1989). Cell 56: 997-1010
- Burtis, K.C., Coschigano, K.T., Baker, B.S. and Wensink, P.C. (1991). EMBOJ 10: 2577-2582

- Büsslinger, M., Schümperli, D. and Birnstiel, M. (1985). Cold Spring Harbor Symp.Quant.Biol. 50: 665-670
- Byrne, G. and Ruddle, F. (1989). Proc.Natl.Acad.Sci. USA 86: 5473-5477
- Calzone, F.J., Britten, R.J. and Davidson, E.H. (1987). Mapping of gene transcripts by nuclease protection assays and cDNA primer extension. In: "Guide to molecular cloning techniques." Berger, S.L. and Kimmel, A.R. (eds.) Academic Press: 611-632
- Campos-Ortega, J.A. (1983). Roux's Arch. Dev. Biol. 192: 317-326
- Casabadan, M., Martinez-Arias, A., Shapira, S. and Chou, J. (1983). Methods Enzymol. 100: 293-308
- Cavener, D. (1987). Nucl.Acids.Res. 15: 1353-1361
- Chapman, K. and Wolfner, M. (1988). Dev.Biol. 126: 195-202
- Cline, T. (1978). Genetics 90: 683-698
- Cline, T. (1984). Genetics 107: 231-237
- Cline, T. (1988). Genetics 119: 829-862
- Cohen, M.B., Maybaum, J. and Sadee, W. (1981). J.Biol.Chem. 256: 8713-8717
- Cohen, M.B. and Sadee, W. (1983). Cancer Res. 43: 1587-1591
- Cole, K., Kandala, J. and Kistler, W. (1986). J.Biol.Chem. 261: 7178-7183
- Cooley, L., Xue, F. and Verheyen, E. (1990). Information listed in Abstracts from *Drosophila* Meeting, Crete, 1990
- Cooney, D.A., Jayaram, H.N., Gebeyehu, G., Betts, C.R., Kelley, J.A., Marquez, V.E. and Johns, D.G. (1982). Biochem.Pharmacol. 31: 2133-2136
- Cooney, D.A., Wilson, Y. and McGee, E. (1983). Anal.Biochem. 130: 339-345
- Cronmiller, C., Schedl, P. and Cline, T. (1988). Genes Dev. 2: 1666-1676
- Cummings, M. and King, R.C. (1969). J.Morph. 128: 427-442
- Dalby, B. and Glover, D.M. (1992). Development in press
- Degelmann, A., Hardy, P.A., Perrimon, N. and Mahowald, A.P. (1986). Dev.Biol. 115: 479-489

- De Lotto, R. and Spierer, P. (1986). *Nature* 323: 688-692
- Di Benedetto, A.J., Lakich, D.M., Krugger, W.D., Belote, J.M., Baker, B.S. and Wolfner, M.F., (1987). *Dev. Biol.* 119: 242-251
- Driever, W. and Nüsslein-Volhard, C. (1988a). *Cell* 54: 83-93
- Driever, W. and Nüsslein-Volhard, C. (1988b). *Cell* 54: 95-104
- Driever, W., Siegel, B., Nüsslein-Volhard, C. (1990). *Development* 109: 811-820
- Driever, W. and Nüsslein-Volhard, C. (1989). *Nature* 337: 138-143
- Duan, D.S. and Sadee, W. (1987). *Cancer Res.* 47: 4047-4051
- Duffy, J. and Gergen, P. (1991). *Genes Dev.* 5: 2176-2187
- Edgar, B. and Schubiger, G. (1986). *Cell* 44: 871-877
- Ephrussi, A., Dickinson, L. and Lehmann, R. (1991). *Cell* 66: 37-50
- Fargnoli, J. and Waring, G.L. (1982). *Devl. Biol.* 92: 306-314
- Feinberg, A.P. and Vogelstein, B. (1983). *Anal. Biochem.* 132: 6-13
- Feinberg, A.P. and Vogelstein, B. (1984). *Anal. Biochem.* 137: 266-267
- Fischer, J., Giniger, E., Maniatis, T. and Ptashne, M. (1988). *Nature* 332: 853-856
- Freeman, M. (1991). *Current Biol.* 1: 378-381
- Frohnhofer, H.G. and Nüsslein-Volhard, C. (1986). *Nature* 324: 120-125
- Frohnhofer, H.G. and Nüsslein-Volhard, C. (1987). *Genes and Development* 1: 880-890
- Gans, M., Audit, C. and Masson, M. (1975). *Genetics* 81: 683-704
- Garabedian, M., Hung, M. and Wensink, P. (1985). *Proc.Natl.Acad.Sci. USA* 82: 1396-1400
- Garabedian, M.J., Shepherd, B.M. and Wensink, P.C. (1986). *Cell* 45: 859-867
- Garcia-Bellido, A. (1964). *Z.Naturforsch* 19: 491-495

- Garcia-Bellido, A. and Robbins, L. (1983). *Genetics* 103: 235
- Gaul, U. and Jäckle, H. (1987) *Cell* 51: 549-555
- Gavin, J. and Williamson, J. (1976). *J.Insect Physiol.* 22: 1737-1742
- Gehring, W., Wieschaus, E. and Holliger, H. (1976). *J.Embryol.Exp.Morph.* 35: 607-616
- Geremia, R., D'Agostino, A. and Monesi, V..(1978). *Exp.Cell Res.* 111: 23-30
- Gerttula, S., Jin, Y. and Anderson, K.V. (1988) *Genetics* 119: 123-133
- Ghosh, S., Gifford, A., Riviere, L., Tempst, P., Nolan, G. and Baltimore, D. (1990). *Cell* 62: 1019-1029
- Gibson, G. and Gehring, W. (1990). *Drosophila Inf.Serv.* 69 in press (cited in Wilson et.al., 1990).
- Gilbert, H.J., Lowe, C.R. and Drabble, W.T. (1979). *Biochem.J.* 183: 481-494
- Gönczy, P., Viswanathan, S. and Di Nardo, S. (1992). *Development* 114: 89-98
- Gould-Somero, M. and Holland, L. (1974). *Wilhelm Roux Arch.EntwMech.Org.* 174: 133-148
- Griffin-Shea, R., Thireos, G. and Kafatos, F. (1982). *Dev.Biol.* 91: 325-336
- Grossniklaus, U., Bellen, H.J., Wilson, C. and Gehring, W. (1989). *Development* 107: 189-200.
- Gubler, U. and Hoffmann, B. (1983). *Gene* 25: 263-269
- Hackstein, J., Beck, H., Hochstenbach, R., Kremer, H. and Zacharias, H. (1990). *Roux's Arch.Dev.Biol.* 199: 251-280
- Haenlin, M., Roos, C., Cassab, A. and Mohier, E. (1987). *EMBO J.* 6: 801-807
- Hafer, E. and Basler, K. (1991). *Development supplement*: 123-131
- Hall J., Gelbart, W. and Kankel, D. (1976). *Mosaics systems. In: Genetics and Biology of Drosophila.* 1a 265-314 Academic Press. London

- Hardy, R., Tokuyasu, K., Lindsley, D. and Garivito, M. (1979). *J.Ultrastruct.Res.* 69: 180-190
- Hashimoto, C., Hudson, K.L. and Anderson, K.V. (1988). *Cell* 52: 269-279
- Hay, B., Jan, L.Y. and Jan, Y.N. (1988). *Cell* 55: 577-587
- Hazelrigg, T. (1987). *Trends in Genetics* 3: 43-47
- Hazelrigg, T., Watkins, W., Marcey, D., Tu, C., Karow, M. and Lin, X. (1990). *Genetics* 126: 607-617
- Hedley, M. and Maniatis, T. (1991). *Cell* 65: 579-586
- Hemmati-Brivanlou, A., Frank, D., Bolce, M., Brown, B., Sive, H. and Harland, R. (1990). *Development* 110: 325- 330
- Hess, O. (1980). Lampbrush Chromosomes. In: *The Genetics and Biology of Drosophila*. Ashburner, M. and Wright, T. (eds.) Vol. 2d: 1-30
- Hess, O. and Meyer, G. (1968). *Adv.Genet.* 14: 171-223
- Hilfiker, A. and Nöthiger, R. (1991). *Roux's Arch.Dev.Biol.* 200: 240-248
- Hoshijima, K., Inoue, K., Higuchi, I., Sakamoto, H. and Shimura, Y. (1991). *Science* 252: 833-836
- Huynh, T., Young, R. and Davis, R.W. (1985). Constructing and Screening cDNA libraries in λ gt10 and λ gt11. In: *DNA Cloning Techniques: A Practical Approach*. D. Glover (ed). IRL Press, Oxford pp. 49-78.
- Illmensee, K. and Mahowald, A.P. (1974). *Proc.Natl.Acad.Sci. USA* 71: 1016-1020
- Illmensee, K. and Mahowald, A.P. (1976). *Exp. Cell Res.* 97: 127-140
- Ingham, P.W. (1988). *Nature* 335: 25-34
- Jäckle, H., Tautz, D., Schuh, R., Seifert, E. and Lehmann, R. (1986) *Nature* 324: 668-670
- Jackson, R.C. and Weber, G. (1975). *Nature* 256: 331-333
- Jackson, R.C., Morris, H.P. and Weber, G. (1977). *Biochem.J.* 166: 1-10
- Jackson, R.J. and Standart, N. (1990). *Cell* 62: 15-24

- Jayaram, H.N., Dion, R.L., Glazer, R.I., Johns, D.G., Robins, R.K., Srivastava, P.C. and Cooney, D.A. (1982). *Biochem.Pharmacol.* 31: 2371-2380
- Jendrisak, J., Young, R.A. and Engel, J.D. (1987). *Methods Enzymol.* 152: 359-371
- Kaiser, K. and Goodwin, S. (1990). *Proc.Natl.Acad.Sci. USA* 87: 1686-1690
- Kalfayan, L. and Wensink, P. (1982). *Cell* 29: 91-98
- Kaltschmidt, B., Glätzer, K., Michiels, F. and Renkawitz-Pohl, R. (1991). *Eur.J.Cell Biol.* 54: 110-120
- Kambyssellis, M.P. (1968). *Univ.Texas Publ.* 6818: 93-134
- Kambyssellis, M. and Gelti-Douka, H. (1974). *Genetics* 77 s33.
- Karn, Brenner, S., Barnett, L. and Cesareni, G. (1980). *Proc.Natl.Acad.Sci. USA* 77: 5172
- Kemphues, K., Kaufmann, T., Raff, R. and Raff, E. (1982). *Cell* 31: 655-670
- Kieran, M., Blank, V., Logeat, F., Vandekerckhove, J., Lottspeich, F., LeBail, O., Urban, M., Kourilsky, P., Baeuerle, P. and Israel, A. (1990). *Cell* 62: 1007-1018
- Kim-Ha, J., Smith, J.L and Macdonald, P. (1991) *Cell* 66: 23-35
- King, R.C., Robinson, A. and Smith, R. (1956). *Growth* 20: 121-157
- Kinzler, K. and Vogelstein, B. (1989). *Nucl.Acids Res.* 10: 3645-3653
- Klingler, M., Erdélyi, M., Szabad, J. and Nüsslein-Volhard, C. (1988). *Nature* 335: 275-277
- Krawczyk, Z., Wisniewski, J. and Biesiada, E. (1987). *Mol.Biol.Rep.* 12: 27-34
- Kuhn, R., Schäfer, U. and Schäfer, M. (1988). *EMBOJ* 7: 447-454
- Kunes, S. and Steller, H. (1991). *Genes Dev.* 5: 970-983
- Lasko, P.F. and Ashburner, M. (1988). *Nature* 335: 611-617
- Lasko, P.F. and Ashburner, M. (1990). *Genes and Development* 4: 905-921

- Laugé, G. (1982). Development of The Genitalia and Analia.
In: *A Handbook of Drosophila Development*. R.Ransom
(ed.) Elsevier Biomedical. pp 237-261
- Lehmann, R. (1988). Development 104 (Supplement): 17-27
- Lehmann, R. and Nüsslein-Volhard, C. (1986). Cell 47:
141-152
- Lehmann, R. and Nüsslein-Volhard, C. (1987) Nature 329:
167-170
- Lehmann, R. and Nüsslein-Volhard. (1991) Development 112:
679-693
- Liddell, S. and Bownes, M. (1991). Mol.Gen.Genet. 230:
219-224
- Lifschytz, E. and Hareven, D. (1977). Dev.Biol. 58: 276-294
- Lindsley, D. and Grell, E. (1968). *Genetic Variations of
Drosophila melanogaster*. Carnegie Institute of
Washington. Publication no. 627
- Lindsley, D. and Tokuyasu, K. (1980). Spermatogenesis. In:
The Genetics and Biology of Drosophila. Vol. 2d: 226-
287. Ashburner, M. and Wright, T. (eds). Academic
Press.
- Littlefield, C. and Bryant, P. (1979). Dev.Biol. 70:
127-148
- Logan, S., Garabedian, M. and Wensink, P. (1989). Genes
Dev. 3: 1453-1461
- MacDonald, P.M. and Struhl, G. (1988). Nature 336: 595-598
- Mahowald, A.P. and Kambyzellis, M.P. (1980). Oogenesis.
In: *The Genetics and Biology of Drosophila*, Vol. 2d,
Ashburner, M., and Wright, T.R.F. eds. Academic Press, New
York, pp. 141-224
- Mahowald, A.P. (1971). J.Exp.Zool. 176: 345-352
- Manseau, L.J. and Schüpbach, T. (1989). Trends in Genetics
5: 400-405
- Marcus, C.H. (1985). *Drosophila* Information Service Volume
61: 193-194
- Mariani, B.D., Lingappa, J.R. and Kafatos, F.C. (1988).
Proc. Natl. Acad. Sci. USA 85: 3029-3033
- Marsh, J. and Wieschaus, E. (1978). Nature 272: 249-251

- Mattox, W. and Baker, B. (1991). Genes and Development 5: 786-796
- Mattox, W., Palmer, M. and Baker, B. (1990). Genes Dev. 4: 789-805
- McKearin, D. and Spradling, A. (1990). Genes Dev. 4: 2242-2251
- McKeown, M., Belote, J. and Boggs, R. (1988). Cell 53: 887-895
- Melton, D.A., Kreig, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984). Nucl. Acids Res. 12: 7035-7056
- Michiels, F., Gasch, A., Kaltschmidt, B. and Renkawitz-Pohl, R. (1989). EMBOJ 8: 1559-1565
- Michiels, F., Wolk, A. and Renkawitz-Pohl, R. (1991). Nucl. Acids Res. 19: 4515-4521
- Mohler, J.D. (1977). Genetics 85: 259-272
- Mohler, J.D. and Wieschaus, E.F. (1986). Genetics 112: 803-822
- Monsma, S.A. and Wolfner, M.F. (1988). Genes and Development 2: 1063-1073
- Mulligan, P., Mohler, J. and Kalfayan, L. (1988). Mol. Cell. Biol. 8: 1481-1488
- Murre, C., McCaw, P. and Baltimore, D. (1989). Cell 56: 777-783
- Nagoshi, R., McKeown, M., Burtis, K., Belote, J. and Baker, S. (1988). Cell 53: 229-236
- Natsumeda, Y., Ikegami, T., Murayama, K. and Weber, G. (1988). Cancer Res. 48: 507-511
- Natsumeda, Y., Ohno, S., Kawasaki, H., Konno, Y., Weber, G. and Suzuki, K. (1990). J. Biochem. 265: 5292-5295
- Nöthiger, R., Dübendorfer, A. and Epper, F. (1977). Wilhelm Roux Arch. Entwicklungsmech. Org. 181: 367-373
- Nöthiger, R., Jonglez, M., Leuthold, M., Meier-Gerschwiler, P. and Weber, T. (1989). Development 107: 505-518
- Nüsslein-Volhard, C. (1990). Drosophila Information Service Volume 68
- Nüsslein-Volhard, C. (1991). Development supplement: 1-11

- Nüsslein-Volhard, C., Wieschaus, E. and Kluding, H. (1984).
 Roux's. Arch. Dev. Biol. 193: 267-282
- Nüsslein-Volhard, C., Frohnhofer, H.G. and Lehmann, R.
 (1987). Science 238: 1675-1681
- Okada, M., Shimura, K., Shiraki, H. and Nakagawa, H.
 (1983). J.Biochem. (Tokyo) 94: 1605-1613
- Olah, E., Natsumeda, Y., Ikegami, T., Kote, Z., Horanyi,
 M., Szelenyi, J., Paulik, E., Kremmer, T., Hollan, S.,
 Sugar, J. and Weber, G. (1988). Proc.Nat.Acad.Sci. USA
85: 6533-6537
- Oliver, B., Perrimon, N. and Mahowald, A. (1988). Genetics
120: 159-171
- Oliver, B., Perrimon, N. and Mahowald, A. (1990). Genetics
125: 535-550
- O'Kane, C.J. and Gehring, W.J. (1987). Proc. Natl. Acad.
 Sci. USA 84: 9123-9127
- Pantelouris, E. (1955). J.Embryol.Exp.Morphol. 3: 86-91
- Parkhurst, S., Bopp, D. and Isch-Horowicz, D. (1990). cell
63: 1179-1191
- Parks, S. and Spradling, A. (1987). Genes.Dev. 1: 497-509
- Pauli, D. and Mahowald, A.P. (1990). Trends in Genetics 6:
 259-264
- Perrimon, N., Engstrom, L. and Mahowald, A. (1984).
 Dev.Biol. 105: 404-414
- Perrimon, N., Mohler, D., Engstrom, L. and Mahowald. A.P.
 (1986). Genetics 113: 695-712
- Pirrota, V. (1986). Cloning *Drosophila* Genes.
 In: " *Drosophila* A Practical Approach." D.B. Roberts
 (ed.) IRL Press: 83-109
- Postlethwaite, J., Bownes, B. and Jowett, T. (1980).
 Dev.Biol. 79: 379-387
- Postlethwaite, J. and Handler, A. (1978). Dev.Biol. 67:
 202-213
- Proffit, R.T., Pathak, V.K., Villa corte, D.G. and Presant,
 C.A. (1983). Cancer Res. 43: 1620-1623
- Prost, E., Deryckere, F., Roos, C. Haenlin, M., Pantesco,
 V. and Mohier, E. (1988). Genes Dev. 2: 891-900

- Raff, J.W., Whitfield, W.G.F. and Glover, D.M. (1990).
Development 110: 1249-1261
- Robertson, H.M., Preston, C.R., Phillis, R.W., Johnson-Schlitz, D.M., Benz, W.K. and Engels, W.R. (1988).
Genetics 118: 461-470
- Romano, C., Bienz-Tadmor, B., Mariani, B. and Kafatos, F. (1988). EMBOJ 7: 783-790
- St. Johnston, R.D. and Gelbart, W.M. (1987). EMBOJ 6:
2875-2791
- St. Johnston, D., Driever, W., Berleth, T., Richstein, S.,
Nüsslein-Volhard, C. (1989) Development supplement
13-19
- St. Johnston, D., Beuchle, D., Nüsslein-Volhard, C. (1991)
Cell 66: 51-63
- Sander, K. and Lehmann, R. (1988). Nature 335: 68-70
- Schäfer, M., Kuhn, R., Bosse, F. and Schäfer, U. (1990).
EMBOJ 9: 4519-4525
- Schäfer, U. (1986). Mol. Gen. Genet. 202: 219-225
- Schulz, R.A. and Butler, B.A. (1989). Genes and
Development 3: 232-243
- Schupbach, T. (1982). Dev. Biol. 89: 117-127
- Schupbach, T. (1987). Cell. 49: 699-707
- Schupbach, T. and Wieschaus, E. (1986) Roux's Arch. Dev.
Biol. 195: 302-317
- Schupbach, T. and Wieschaus, E., (1989). Genetics 121:
101-117
- Schupbach, T., Wieschaus, E. and Nöthiger, R. (1978).
Wilhelm Roux Arch. Entwicklungsmech. Org. 181: 367-373
- Searles, L., Jokerst, R., Bingham, P., Voelker, R. and
Greenleaf, A. (1983). Cell 31: 585-592
- Shirras, A.D. and Bownes, M. (1987). Mol. Gen. Genet. 210:
153-155
- Singh, H., LeBowitz, J., Baldwin, A. and Sharp, P. (1988).
Cell 52: 415-423
- Slee, R. and Bownes, M. (1990). Quarterly Review of
Biology 65: 175-204

- Sonnenblick, B. (1941). Proc.Natl.Acad.Sci. USA 27:
484-489
- Sonnenblick, B. (1950). In: *The Biology of Drosophila*.
(M.Demerec, ed.). John Wiley and Sons, New York.
pp.62-167
- Sosnowski, B., Belote, J. and McKeown, M. (1989). Cell 58:
449-459
- Southern, E. (1975). J.Mol.Biol. 98: 503-517
- Spradling, A.C. and Mahowald, A.P. (1980). Proc. Natl.
Acad. Sci. USA 77: 1096-1100
- Spradling, A.C., Waring, G.L. and Mahowald, A.P.(1979).
Cell 16: 609-616
- Sprenger, F., Stevens, L.M. and Nüsslein-Volhard, C.
(1989). Nature 338: 478-483
- Stein, D., Roth, S., Vogelsang, E. and Nüsslein-Volhard, C.
(1991). Cell 65: 725-735
- Steinhauer, W., Walsh, R. and Kalfayan, L. (1989).
Mol.Cell.Biol. 9: 5726-5732
- Steinmann-Zwicky, M. (1988). EMBOJ 7: 3889-3898
- Steinmann-Zwicky, M., Fuhrer, E., Franken, D. and
Nöthiger, R. (1987). Information reported at 10th
European *Drosophila* conference in Barcelona.
- Steinmann-Zwicky, M., Schmid, H. and Nöthiger, R. (1989).
Cell 57: 157-166
- Stern, C. and Hadorn, E. (1938). Amer.Nat. 72: 42-52
- Stevens, L., Frohnhofer, H., Klingler, M. and Nüsslein-
Volhard, C. (1990). Nature 346: 660-663
- Steward, R. (1987). Science 238: 692-694
- Steward, R., Zusman, S., Huang, L. and Schedl, P. (1988).
Cell 55: 487-495
- Strecker, T.A., Merriam, J.R. and Lengyel, J.A. (1988).
Development 102: 721-734
- Suter, B., Romberg, L.M. and Steward, R. (1989). Genes and
Development 3: 1957-1968
- Tautz, D. et.al (1987). Nature 327: 383-389
- Tautz, D. (1988). Nature 332: 281-284
- Tautz, D. and Pfeifle, C. (1989). Chromosoma 98: 81-85

- Thisse, B., Stoetzel, C., El Messal, M. and Perrin-Schmitt, F. (1987). *Genes Dev.* 1: 709-715
- Thomas, P.S. (1980). *Proc.Natl.Acad.Sci. USA* 77: 5201-5205
- Tiedeman, A.A. and Smith, J.M. (1985). *Nucleic Acids Res.* 13: 1303-1316
- Tiedeman, A.A. and Smith, J.M. (1991). *Gene* 97: 289-293
- Torres, M. and Sanchez, L. (1989). *EMBOJ* 8: 3079-3086
- Tricot, G.J., Jayaram, H.N., Lapis, E., Natsumeda, Y., Nichols, C.R., Kneebone, P., Heerama, N., Weber, G. and Hoffman, R. (1989). *Cancer Res.* 49: 3696-3701
- Underwood, E., Caulton, J. and Mahowald, A.P. (1979). *J.Cell.Biol* 83: 225a
- Van Deusen E. (1976). *J.Embryol.Exp.Morphol.* 37: 173-185
- Wakimoto, B.T., Turner, F.R. and Kauffmann, T.C. (1984). *Dev. Biol.* 102: 147-172
- Wang, C. and Lehmann, R. (1991) *Cell* 66: 637-647
- Watanabe, T. (1975). *Jpn.J.Genet.* 50: 269-271
- Weber, G. (1983). *Cancer.Res.* 43: 3466-3492
- Weber, G., Jayaram, H.N., Lapis, E., Natsumeda, Y., Yamada, Y., Yamaji, Y., Tricot, G.J. and Hoffmann, R. (1988). *Adv.Enzyme Regul.* 27: 405-433
- Weber, G., Olah, E., Lui, M., Kizaki, H., Tzeng, D. and Takeda, E. (1983). *Adv.Enzyme Regul.* 18: 3-26
- Wharton, R.P. and Struhl, G. (1989). *Cell* 59: 881-892
- Wharton, R.P. and Struhl, G. (1990). Information listed in Abstracts from *Drosophila Meeting, Crete, 1990*
- Whitfield, W.G.F., Gonzalez, C., Sanchez-Herrero, H. and Glover, D.M. (1989). *Nature* 338: 337-340
- Wieschaus, E. (1979). In *Cell lineage, Stem cells and Cell Differentiation*. LeDourain, N. (ed): 291-302
- Wieschaus, E., Audit, C. and Masson, M. (1981). *Devl. Biol.* 88: 92-103
- Wieschaus, E., Marsh, J.L. and Gehring, W. (1978). *Wilhelm Roux's Arch. Dev. Biol.* 184: 75-82
- Wieschaus, E. and Szabad (1979). *Devl. Biol.* 68: 29-46

- Wilson C., Bellen, H. and Gehring, W. (1990).
Annu.Rev.Cell.Biol. 6: 679-714
- Wilson, C., Pearson, R.K., Bellen, H., O'Kane, C.J.,
Grossniklaus, U. and Gehring, W. (1989). Genes and
Development 3: 1301-1313
- Wilson, K., Collart, F.R., Huberman, E., Stringer, J. and
Ullman, B. (1991). J.Biochem. 266: 1665-1671
- Wong, Y., Pustell, J., Spoerel, N. and Kafatos, F. (1985).
Chromosoma 92: 124-135
- Yamada, Y., Natsumeda, Y. and Weber, G. (1988).
Biochemistry 27: 2193-2196
- Yanicostas, C., Vincent, A. and Lepesant, J.A. (1989).
Mol. and Cell. Biol. 9: 2526-2535
- Yarger, R.J. and King, R.C. (1971). Dev.Biol. 24: 166-177
- Young, R.A. and Davis, R.W. (1983). Proc.Natl.Acad.Sci.
USA 80: 1194-1198

SEX DETERMINATION IN *DROSOPHILA MELANOGASTER*

by

ROGER SLEE and MARY BOWNES

Reprinted From: Quarterly Review of Biology
Volume 65 Number 2 June 1990

SEX DETERMINATION IN *DROSOPHILA MELANOGASTER*ROGER SLEE
MARY BOWNES*Department of Molecular Biology, University of Edinburgh,
Mayfield Road, Edinburgh EH9 3JR, UK*

ABSTRACT

The understanding of sex determination is a fundamental goal in the study of eukaryotic developmental genetics. The mechanisms governing the generation of sexual dimorphism have been well characterized in *Drosophila* because of its amenability to both genetic manipulation and the application of the techniques of modern molecular genetics.

By using classical genetics to search for sex-transforming mutations and by analysing their phenotypes and how they interact, a picture has emerged involving a cascade of regulatory genes. The primary sex determining signal—the ratio of the number of *X* chromosomes to the number of sets of autosomes—sets this cascade into motion. Genetic evidence has suggested that the intervening genes in this pathway are active in females but not in males, whereas the final gene has active but opposing roles in the two sexes. This bifunctional locus is responsible for the repression of female differentiation genes in males and male differentiation genes in females.

The cloning of the key genes of the regulatory cascade and the study of their transcription patterns have revealed that their different functional states in the two sexes do not result from control at the transcriptional level, as might have been expected. Instead, common primary transcripts are produced in male and female flies; these are then differentially spliced to encode sex-specific gene products.

In this paper we focus on the contributions of molecular genetics to the understanding of sex determination. Sufficient background is included for the reader to see how the models of the *Drosophila* sex determination system were first developed. We then show how the application of new technology has complemented the genetic approach and refined our understanding of the system.

Current intensive research in this area should lead within the next few years to definitive knowledge at the molecular level of the cascade of differential splicing of regulatory genes, and how this hierarchy ultimately gives rise to the appropriate sex-specific patterns of structural gene expression that underlie sexual dimorphism.

INTRODUCTION

SEXUAL DIMORPHISM in *Drosophila melanogaster* adults is extensive. There are morphological differences in the bristle pattern, pigmentation and segmentation of the external cuticle and in the internal reproductive system. Innate behavioral differences between the sexes, reflecting dimorphism in the structure and function of at least part of the fly's nervous system, are also observed. In this paper we review current knowledge of the system responsible for sex determination and differences between the sexes in *Drosophila melanogaster*,

knowledge that has gained from a synthesis of genetic and molecular experimental approaches.

Setting the Primary Signal—The X:A Ratio

It is apparent from observation of aneuploid individuals that the primary determinant of sexual differentiation is the ratio of the number of *X* chromosomes (*X*) to the number of sets of autosomes (*A*) (Bridges, 1921, 1925a,b). Females possess two *X* chromosomes and males have one *X* chromosome and one *Y* chromosome. Individuals with an *X:A* ratio of ≤ 0.5 are male, whereas those with a ratio of ≥ 1 are

female. The Y chromosome has no role in somatic sexual differentiation, but is required in the male for normal spermatogenesis. Flies with an intermediate ratio (e.g., 2X:3A) are intersexes, which are coarse-grained mosaics composed of patches of male and female tissue (Hannah-Alava and Stern, 1957; Stern, 1966; Cline, 1984). This phenotype reflects a disturbance of sexual pathway choice at the level of individual cells that is due to the ambiguity of the primary signal, and demonstrates the cell-autonomous nature of the sexual differentiation process.

The Genetic Cascade or Hierarchy

The primary signal of the X:A ratio serves to determine the functional state of the key regulatory gene *Sex-lethal* (*Sxl*) (Cline, 1978, 1979b). In female flies, *Sxl* is active; this functional state is translated by means of the action of a hierarchical series of autosomal loci into the specific pattern of structural gene expression that underlies female differentiation. Four genes of this autosomal hierarchy have

been genetically characterized thus far: *transformer* (*tra*), *transformer-2* (*tra-2*), *doublesex* (*dsx*) and *intersex* (*ix*), and their respective positions and functions within the hierarchy have been established (Table 1). Mutational analysis suggests that additional genes may be involved in transducing the female signal provided by the active *Sxl* gene, but their characterization remains at a preliminary stage; they are considered only briefly in a later section. In male flies the *Sxl* gene is not activated. As a consequence the genes of the autosomal hierarchy are either inactive or assume alternative functions, with the result that male differentiation genes are expressed.

With the cloning of several of the genes of the regulatory hierarchy, their molecular analysis has reached an advanced stage and it has been possible to confirm many of the interactions inferred earlier from genetic data. It appears that the alternative functional states of these genes in males and females are achieved through differential processing of primary transcription products. Although there are many

TABLE 1
The most studied genes of the regulatory hierarchy of sex determination

Gene	Mutation(s)	Phenotype	Function	Map position
<i>Sex-lethal</i> (<i>Sxl</i>)	Recessive	XX cells → ♂ phenotype; lethal to XX flies	Active in ♀ to determine somatic differentiation by activation of <i>tra</i> and prevent hypertranscription of X chromosome; inactive in ♂	1-19.2
	Dominant	XO cells → ♀ phenotype; lethal to XY flies		
<i>transformer</i> (<i>tra</i>)	Recessive	XX flies → pseudomales	Active in conjunction with <i>tra-2</i> in regulation of <i>dsx</i> in ♀; inactive in ♂	3-4.5
<i>transformer-2</i> (<i>tra-2</i>)	Recessive	XX flies → pseudomales	Active in ♀ to induce ♀-specific <i>dsx</i> expression and repress ♂-specific <i>dsx</i> expression; inactive in ♂ sex determination	2-70
<i>doublesex</i> (<i>dsx</i>)	Recessive	{ XX flies → intersex or XY flies → intersex or XX and XY flies → intersex	Active in ♂ to repress ♀ differentiation functions	84E 1-2
	3 alleles			
	Dominant	XX flies → intersex	Active in ♀ to repress ♂ differentiation functions	
<i>intersex</i> (<i>ix</i>)	Recessive	XX flies → intersex	Active in ♀ in conjunction with <i>dsx</i> product to repress ♂ differentiation functions	2-60.5

known instances of the generation of different cell type-specific proteins from a single gene by this mechanism, such as the *Drosophila* locus *Shaker* (Sutcliffe and Milner, 1988) and the rat calcitonin gene (Amara et al., 1982), this method of controlling regulatory gene activity in sex determination is a new finding; its detailed characterization is helping to elucidate the molecular interactions involved. Moreover, it seems particularly remarkable that in the case of *Sxl* and *tra*, which function only in female flies, this ON/OFF regulation should be at the post-transcriptional level.

In the following sections the genetic and molecular data pertaining to each of the genes of the sex determination hierarchy are discussed in detail. We then consider the differentiation genes at the end of the pathway, and discuss the difficulties inherent in isolating these genes. Finally, we look at the future potential of the system for understanding sex determination in *Drosophila*.

THE HEAD OF THE HIERARCHY

Sex-Lethal

Genetic (Baker and Ridge, 1980) and molecular (McKeown et al., 1988; Nagoshi et al., 1988) studies suggest that the primary signal (X:A ratio) is translated by means of a cascade of regulatory genes into sex-differential expression of the structural genes underlying sexual dimorphism. Of the regulatory genes identified (by mutation) to date, the X-linked gene *Sex-lethal* (*Sxl*) appears to play a pivotal role. Studies suggest that *Sxl* is active only in females; positioned at the head of the hierarchy, it dictates both the female pathway of differentiation and a reduced transcription rate of the two female X chromosomes, relative to that of the single X chromosome of the male (Cline, 1978, 1979a,b; Lucchesi and Skripsky, 1981; Sanchez and Nöthiger, 1982; Gergen, 1987). Diplo-X flies homozygous for *Sxl*^{M#1} die as embryos, but chromosomal males hemizygous for this loss-of-function allele, or with the *Sxl* region deleted, are viable and fertile. Conversely, the gain-of-function mutation *Sxl*^{M#1} is lethal in chromosomal males, but does not affect chromosomal females (Cline, 1978; Marshall and Whittle, 1978) (Table 1). Since chromosomal females transformed into phenotypic males by mutations at either of two other regulatory loci, *transformer* (*tra*) or *doublesex* (*dss*), still require *Sxl*,

and since the presence of a Y chromosome does not protect XX flies from the effect of *Sxl*^{M#1}, the lethality is evidently a consequence of the presence of two X chromosomes (Muller and Zimmering, 1960; Marshall and Whittle, 1978). This interpretation is consistent with the gene having an active role in female dosage compensation, to be discussed later.

The sex-transforming effects of *Sxl* mutations are conveniently observed in mutant clones, an approach that circumvents the problems of zygotic lethality (Table 1). Visibly marked clones homozygous for a loss-of-function *Sxl* mutation can be generated in XX larvae heterozygous for the mutation by radiation-induced mitotic recombination; in sexually dimorphic regions of the adult, such clones undergo male differentiation. In addition, such clones are smaller than controls that carry a wild type *Sxl* gene, as might be expected if disturbed dosage compensation reduced the ability of the mutant clones to compete with neighboring cells (Cline, 1979a; Sanchez and Nöthiger, 1982).

In gynandromorphs, which are mosaics of XX and XO cells resulting from sporadic loss of an unstable ring-X chromosome during syncytial cleavage, the presence of the gain-of-function mutation *Sxl*^{M#1} causes XO cells to undergo female differentiation. These mutant gynandromorphs exhibit much lower viability than wild type controls (Cline, 1979b), probably as a result of incorrect dosage compensation in the XO cells. The reciprocal effects of the two classes of *Sxl* mutations, observed in these studies, demonstrate that an active *Sxl* gene dictates female differentiation and an inactive *Sxl* gene dictates male differentiation.

Experiments involving the removal of one X chromosome at various developmental stages from cells of XX embryos have further elucidated the role of this key gene. When the loss of an X chromosome occurs during syncytial cleavage, before cellularization, the resulting animal is a sexual mosaic (gynandromorph) in which each cell's rate of X chromosome transcription and sexual differentiation autonomously reflects the X:A ratio (Sanchez and Nöthiger, 1983). Removal of an X chromosome at the blastoderm stage, after cell formation, produces clones of poor viability which are gradually eliminated, unless the remaining X carries the loss-of-function mutation *Sxl*^{M#1}, in

which case they survive to form male structures. The conferral of viability on the latter cells by *Sxl*^{#1} is best explained by the proposed role of this gene in dosage compensation: the single X carrying *Sxl*⁺ is transcribed at the female rate, whereas the X chromosome carrying *Sxl*^{#1} is properly dosage compensated. Further evidence for the action of *Sxl* in dosage compensation has come from estimation of X chromosome transcription rates in chromosomally female larvae carrying various hypomorphic *Sxl* mutations; these studies revealed increased transcription rates in the two most extreme mutant combinations (Lucchesi and Skripsky, 1981) and from the effect of the loss of *Sxl* activity on the phenotype of female larvae hemizygous for a hypomorphic allele of the X-linked segmentation gene *runt*. With a wild type *Sxl* gene, these larvae display a mutant phenotype of markedly greater severity than male larvae hemizygous for the same *runt* allele, as expected from dosage compensation considerations. If the female larvae are also homozygous for the loss-of-function allele *Sxl*^{#1}, however, the severity of the *runt* phenotype is reduced to the level of the male, indicating hyperactivation of X-linked genes in the absence of *Sxl* function (Gergen, 1987).

From the preceding data, it would seem that *Sxl* somehow assesses the X:A ratio about the time of the blastoderm stage, and then becomes independent of that signal. The rate of transcription of the X chromosome and the pathway of sexual differentiation followed are, however, continuously dependent on the ON/OFF state of *Sxl* (Sanchez and Nöthiger, 1983). The idea that cells make an early commitment to sexual phenotype, which is maintained in a clonal fashion thereafter, is supported by the size of the patches of male and female tissue observed in triploid intersexes (2X:3A) and by the similarity between the mosaic phenotypes of these flies and gynandromorphs, which are known to result from genetic mosaicism well before blastoderm stage (Cline, 1984). Further evidence that the state of activity of *Sxl* is determined early and irreversibly has been obtained with XX mosaic intersexes resulting from genetic manipulations that lower the probability of *Sxl* activation (Cline, 1985). Visibly marked cell clones in sexually dimorphic regions of the foreleg of these flies were generated by radiation-induced somatic recombi-

nation between the mid-first and early second larval instars. The marked clones always differentiated as either male or female, depending on whether their *Sxl* gene had been activated, and never included tissue of both sexual phenotypes, implying irreversible determination of the state of activity of *Sxl* prior to the time of irradiation (Cline, 1985). The conclusion from these various studies, that *Sxl* activation is irreversible, points to an autoregulatory role for the gene, which maintains its active state following X:A ratio assessment. This notion is supported by the isolation of a double mutant allele, *Sxl*^{fm#7,M#1}, which almost totally lacks sex-determining function, but seems to express constitutively an autoregulatory function, since it is able to activate a wild type *Sxl* gene in *trans* in the absence of the maternally supplied *daughterless* gene product which is normally required for *Sxl* activation, as discussed below (Cline, 1984).

One speculative model for the initial reading of the X:A ratio posits three main elements:

- (1) An autosomally encoded factor present in equal amounts in both sexes which acts as a repressor of *Sxl*.
- (2) Several sites on the X chromosome that bind the repressor with high affinity.
- (3) A *Sxl*-associated binding site of low affinity.

According to this model, the availability of free repressor, and hence the state of activity of *Sxl*, will depend on the X:A ratio (Gadagkar et al., 1982; Chandra, 1985).

Maternal Involvement in Sex-Lethal Activation

Genetic studies indicate that in addition to a female sex-chromosome constitution, activation of *Sxl* requires the products of two maternally acting genes, *daughterless* (*da*) and *Daughter killer* (*Dk*) (Table 2).

At 25° C mothers homozygous for a temperature sensitive *da* allele produce no female progeny, regardless of the paternal genotype at this locus (Bell, 1954). Chromosomally female offspring die prior to adulthood as a result of incorrect dosage compensation (Cline, 1983, 1984). However, daughters of *da* mothers are efficiently rescued by possession of the constitutive *Sex-lethal* mutation *Sxl*^{M#1}. These observations have led to the hypothesis that in female zygotes the X:A ratio, together with the maternally supplied *da* gene product, activates

TABLE 2

Genes that interact with the key regulatory gene Sex-lethal in dosage compensation and sex-determination

Gene	Mutations	Phenotype	Function in sex determination	Map position
<i>daughterless (da)</i>	Recessive: maternally acting	♀ embryonic lethal	Necessary for activation of zygotic <i>Sxl</i> in ♀ embryos	2-41.5
<i>Daughter killer (Dk)</i>	Haploid insufficient: maternally acting	♀ embryonic lethal	Necessary for activation of zygotic <i>Sxl</i> in ♀ embryos	3-34.2
<i>fs(1) 1621</i>	Recessive	Ovarian tumors in ♀; sterile	Required for <i>Sxl</i> activation in ♀ embryos	4F7-4F11
<i>sisterless a and b (sis-a and sis-b)</i>	Recessive	♀ embryonic lethal	Possible X chromosome numerator elements involved in <i>Sxl</i> activation in ♀	<i>sis-a</i> 1-34.3 <i>sis-b</i> 1-0.0
<i>virilizer (vir)</i>	Recessive (multiple alleles)	(a) ♀ and ♂ lethal (b) XX cells → ♂ phenotype (c) XX fly → intersex	ON in ♀ to translate active state of <i>Sxl</i> into ♀ somatic differentiation and basal level of X chromosome transcription (possible additional vital nonsex-specific function)	2-103.3
<i>male-specific lethals 1, 2, 3 (msl 1, 2, 3)</i> <i>maleless (mle)</i>	Recessive	♂ lethal	Translation of inactive state of <i>Sxl</i> into X chromosome hypertranscription in ♂	<i>msl 1</i> 2-53.3 <i>msl 2</i> 2-9.0 <i>msl 3</i> 3-25.8 <i>mle</i> 2-56.8
<i>hermaphrodite (her)</i>	Recessive	Feminizes XY flies; masculinizes XX flies; reduces viability of both sexes	Required for normal sex-determination in both sexes (possible additional vital nonsex-specific function)	2-52.9

Sxl, and that the female-specific lethality of the *da* mutation is a consequence of hypertranscription of the two X chromosomes, in the absence of *Sxl* product. The observation that 2X:3A progeny of *da/da* mothers (which do not show appreciably reduced viability) are much more male-like than the triploid intersex progeny of *da⁺/da⁺* mothers confirms that the *da⁺* product is necessary for the action of *Sxl* in sex determination, as well as in dosage compensation (Cline, 1983). In addition to acting maternally in the activation of *Sxl* in female progeny, *da* also has several zygotic functions (Cline, 1976; 1980). Zygotic expression of the gene is required in both sexes for the formation of all peripheral neuronal and associated sensory structures (Caudy et al., 1988) and the adult cuticle (Cronmiller and Cline, 1987). Also, adult females require somatic *da* expression for proper egg-membrane formation. Thus, after prolonged exposure to 29°C, females homozygous for the temperature sensitive *da* mutation produce in-

fertile eggs (Cline, 1976, 1980; Cronmiller and Cline, 1987).

The *daughterless* locus has been cloned (Cronmiller et al., 1988) using the P-element transposon tagging method (Bingham et al., 1981). The transcription unit is less than 5 kb in length and encodes two nonsex-specific overlapping transcripts of 3.2 and 3.4 kb; they are present at similar levels in all developmental stages except in embryos up to 2.5 hours old, when the smaller transcript appears to be relatively more abundant, making it a candidate for encoding the maternal *da* function. Caudy et al. (1988) have cloned *da* independently, and reported two transcripts of 3.2 kb and 3.7 kb, only the smaller of which is detected in ovarian RNA. The presence of a 3.2 kb *da* transcript in adult males, however, argues against an exclusively maternal role for this transcript (Cronmiller et al., 1988). It is currently uncertain whether the *da* transcripts differ in coding potential.

Although this relatively simple transcription

pattern may seem at odds with the functional complexity of the *da* gene, it may be significant that all the *da* alleles characterized to date seem to affect *all* of the *da* functions, which would be consistent with the idea of a single multifunctional *da* gene product (Cronmiller et al., 1988).

The sequence of the predicted 74 kd protein product of a cloned *da* cDNA shows significant similarity to the *bicoid* and *paired* gene products (Cronmiller et al., 1988), both of which are regulatory proteins involved in *Drosophila* segmentation (Berleth et al., 1988; Frigerio et al., 1986). The *da* protein product also shows similarities to the products of the *myc* and *MyoDI* genes and the *T3*, *T4* and *T5* genes of the *Drosophila achaete-scute* complex (Caudy et al., 1988). Like *da*, the latter five genes are all implicated in processes of cellular differentiation (Bishop, 1985; Davis et al., 1987; Garcia-Bellido, 1979), and in the case of the *achaete-scute* complex the cells concerned are neuronal precursors in the peripheral nervous system. It seems reasonable to suppose, therefore, that these sequence similarities reflect functional similarities between *da* and the various genes mentioned, involving control of expression of differentiation genes. In this regard, it is notable that the *myc* gene product has been implicated in transcriptional regulation (Kingston et al., 1984) and RNA processing (Sullivan et al., 1986). Further support for a *da* role in transcriptional regulation comes from recent reports of a striking similarity between the *da* product and an enhancer binding protein (Murre et al., 1989).

The *Sxl* gene has been cloned (see below), and Northern analysis has shown that the transcription pattern at the blastoderm stage gives way to a second pattern of male- and female-specific transcripts which persists throughout subsequent development (Maine et al., 1986; Bell et al., 1988). Although it is currently unknown whether any of the early transcripts are female specific, it has been proposed that the maternally provided *da* product could be a transcription factor that is necessary for the production of an early female-specific *Sxl* protein at a time when promoters are not generally active (Schüpbach, 1985; Cronmiller et al., 1988). This hypothetical early female-specific *Sxl* product would then act to impose the later female-specific pattern of *Sxl* transcripts that

has been observed (discussed further in a later section); it also provides for the early dosage compensation function of the gene which has been shown to operate as early as the blastoderm stage, and to be dependent on maternal *da* product (Gergen, 1987). In agreement with such a role for *da*, maternal *da* mutations have been reported to alter the early *Sxl* transcription pattern in some undefined way, and high levels of the normally adult male-specific *Sxl* transcripts are detected in surviving chromosomally female progeny of homozygous *da* mothers (Maine et al., 1986).

More recently *Daughter killer* (*Dk*), a second locus whose maternal product is required for *Sxl* activation, has been identified. Like the original *da* mutation, the dominant mutation *Dk* is a temperature-sensitive maternal effect lethal that affects female embryos. Again, the mutant phenotype can be rescued by a zygotic constitutive *Sxl* mutation, and again there is a masculinizing effect on 2X:3A progeny. This overlap in mutant phenotypes suggests that the *da* and *Dk* gene products act together to activate *Sxl* in female embryos. Temperature-shift experiments have revealed that the requirement for functional *Dk* product ceases before the end of oogenesis, whereas the *da* product is required until the blastoderm stage, when the state of *Sxl* is set. Thus, *Dk* may not regulate *Sxl* directly, but may act through an intermediary. The *Dk* alleles examined to date are homozygous lethal to both sexes, suggesting that, like *da*, *Dk* may have different maternal and zygotic functions (M. Steinmann-Zwicky, E. Fuhrer-Bernhardsgrütter, D. Franken, and R. Nöthiger, unpubl.).

Zygotically Acting Genes that Interact with Sex-Lethal

In addition to the maternally acting genes *da* and *Dk*, a number of zygotically acting genes appear to be involved in either the activation of *Sxl* in females and/or the mediation of the gene's functions in somatic sex determination or dosage compensation (Table 2). To date, three loci have been identified which appear to be in the first category. These are *sisterless-a* and *sisterless-b* (*sis-a* and *sis-b*) and *fs(1)1621*.

The gene *sis-a* (1-34.3) was first identified as an X-linked, recessive, zygotically acting, female-specific lethal with a masculinizing effect on triploid intersexes (2X:3A) (Cline,

1986). It was shown subsequently that a duplication of the *sis-a*-containing region of the X chromosome was lethal to chromosomally male flies that also carried two doses of *Sxl* (Cline, 1988). The same study revealed the existence of a second, separate locus on the X chromosome in the region of the *achaete-scute* complex, which interacted synergistically with *sis-a* and was named *sisterless-b* (*sis-b* 1-0.0). Simultaneous duplication of *sis-a* and *sis-b* is lethal to chromosomal males, whereas simultaneous deletion of one copy of both loci is lethal to chromosomal females. These reciprocal effects on male and female viability, together with the interaction with *Sxl*, point to the possibility that *sis-a* and *sis-b* are positive regulators of *Sxl*, acting in a quantitative fashion to communicate the number of X chromosomes to the *Sxl* gene in male and female embryos and to ensure its appropriate state of activity, ON or OFF. As such, these loci have been designated X:A numerator elements (Cline, 1988), members of a class of genetic elements whose relative dose signals the X:A ratio. It remains to be determined whether these loci encode *trans*-acting products or are rather targets for an autosomal repressor of *Sxl*, as envisaged in Chandra's model of *Sxl* regulation (Chandra, 1985). There is evidence, however, that *sis-b* may in fact correspond to the *scute* locus of the *achaete-scute* complex (AS-C) since mutations at this locus behave like a *sis-b* deficiency (Cline, 1988; Torres and Sanchez, 1989). Thus *scute* may have dual functions, in sex determination and neurogenesis, perhaps acting in conjunction with *da*, which is involved in the same processes and has sequence similarity to the T4 product of the AS-C, which corresponds to the active product of the *scute* locus (Campuzano et al., 1985).

The following genetic interactions lend support to the proposed model of *sis-a/sis-b* action (Cline, 1988). The male-specific lethality of simultaneous duplications for *sis-a* and *sis-b* is rescued by a zygotic loss-of-function *Sxl* mutation or a maternal *da/da* genotype. Triploid intersexes (2X:3A) in which the ambiguous X:A ratio results in cell clones of both male and female phenotype are very sensitive to a single extra dose of *sis-b*, which increases the percentage of entirely female sexually dimorphic structures from 0 to 37 percent. In connection with the latter result, it is assumed that the sex-transforming effects of *sis-a/sis-b* gene dosage

in diploid flies are masked by concomitant lethality that is due to disturbed dosage compensation, whereas in triploid flies the ratio between the levels of X chromosomal and autosomal gene expression is less affected by aberrant dosage compensation (Cline, 1988).

A previous study concluded that a different region of the X chromosome (3E8-4F11) contained similar X:A numerator elements, based on the following observations:

(1) In a study of aneuploid cell clones, one complete X chromosome and a distal duplication of the region 1A to 7D (X/D 1A-7D) resulted in female differentiation, whereas cells carrying an X chromosome deleted for the 3E8-4F11 region and the same duplication (Df 3E8-4F11/Dp1A-7D) differentiated as male.

(2) A duplication for this region was dominant male-specific lethal, whereas females or males carrying a loss of function *Sxl* mutation (*Sxl^l*) were unaffected.

(3) Deletion of one copy of the region was lethal or masculinizing in females heterozygous for *Sxl^l* (Steinmann-Zwicky and Nöthiger, 1985).

However, more recent data have shown that the 3E8-4F11 region does not conform to the expectations of an X:A numerator element in several respects. First, the elements within the region responsible for the male and female lethality have been shown to be genetically separable (Cline, 1988). Second, attempts to reproduce the reported rescue of duplication-bearing males by *Sxl^l* have failed (Oliver et al., 1988; Cline, 1988). Third, the viability of females heterozygous for a deletion of 3E8-4F11 and *Sxl^l* is very dependent on the maternal genotype with respect to this region (no maternal effect would be expected of an X:A numerator element), while the degree of masculinization of female escapers of the lethal interaction was independent, indicating that the lethality and masculinization effects reflect distinct genetic interactions (Cline, 1988).

Deficiency mapping has enabled the further localization of the region involved in the female lethal interaction with *Sxl^l* to 4F7-4F11 (Cline, 1988). The gene responsible is the previously identified *fs(1)1621* (Gans et al., 1975; Gollin and King, 1981; Perrimon et al., 1986) since renamed *sans-fille* (Oliver et al., 1988) and *liz* (Steinmann-Zwicky, 1988).

Females heterozygous for a loss-of-function

fs(1)1621 mutation and a loss-of-function *Sxl* mutation exhibit the somatic sex transformation and greatly reduced viability observed by Steinmann-Zwicky and Nöthiger (1985), which is consistent with *fs(1)1621* acting as a positive regulator of *Sxl* in the female zygote (Oliver et al., 1988). Previous studies have shown similar interactions between *Sxl* gene dosage and upstream elements known to be involved in *Sxl* activation (Cline, 1978, 1983, 1988). In addition, these double heterozygotes are usually sterile because of ovarian tumors, a phenotype also seen in germ line clones that lack *Sxl* function (Schüpbach, 1985). The constitutive *Sex-lethal* mutation, *Sxl^{M#1}*, was shown to rescue the lethality and sex transformation of such double heterozygotes, implying that *fs(1)1621* does act upstream of *Sxl* (Steinmann-Zwicky, 1988). However, the same study also showed that the male lethality of *Sxl^{M#1}* was rescued by the loss-of-function *fs(1)1621* mutation, a finding not predicted by such a model of the latter gene's action; the explanation was offered that *Sxl^{M#1}* may not always be active, as proposed (Maine et al., 1986), and is sensitive to upstream signals such as the X:A ratio and positive regulators such as *fs(1)1621* (Steinmann-Zwicky, 1988). Earlier data have also indicated that *Sxl^{M#1}* is not truly constitutive in either dosage compensation or sex determination functions (Cline, 1979b, 1980, 1984; Gergen, 1987).

Somewhat surprisingly, both studies found the viability and somatic sexual phenotype of homozygous mutant *fs(1)1621* females to be unaffected in the presence of two wild type *Sxl* genes. Similar mutations, which only give a mutant phenotype in conjunction with a mutation at a second locus, have previously been identified in *Caenorhabditis elegans* (Ferguson et al., 1987). Homozygous *fs(1)1621* females do, however, exhibit the sterility and ovarian tumor phenotype observed in the double mutants, suggesting that *Sxl* function in the germ line may be more sensitive to perturbation by *fs(1)1621* mutations (Oliver et al., 1988). Steinmann-Zwicky (1988) has proposed that *fs(1)1621* is in fact a hypomorphic allele providing sufficient product for somatic but not for germ line *Sxl* activation.

The observation that maternal mutations at the *fs(1)1621* locus greatly reduce the viability of female progeny heterozygous for a loss-of-

function *Sxl* allele, but do not affect sexual phenotype, suggests that *fs(1)1621* product inherited through the maternal germ line is required for early *Sxl* expression, whereas later expression is dependent on zygotic *fs(1)1621* gene product. According to this model, the absence of a maternal effect on somatic sex reflects the fact that, in flies which escape the lethal maternal effect of loss of *fs(1)1621* function on early *Sxl* expression, sexual phenotype is entirely dependent on zygotic *fs(1)1621* expression (Oliver et al., 1988; Steinmann-Zwicky et al., 1988).

Genes appearing to act after *Sxl* in the dosage compensation pathway are the recessive *male-specific lethals* (*msl 1, 2, 3*) and *maleless* (*mle*), which kill males exclusively throughout larval life. The observation that triple homozygous mutants for *msl 1* and *2* and *mle* survive as long as single mutants at these loci (Belote, 1983) suggests that these three genes act in the same pathway. These genes are candidates for translating the OFF state of *Sxl* into hyperactivation of the single X chromosome in male flies, probably by interaction with *cis*-acting sequences associated with X-linked genes (Pirrotta et al., 1985; Krumm et al., 1985). The following lines of evidence, however, indicate that additional genes must also be involved, at least in early dosage compensation functions:

(1) Mutations at these male-specific lethal loci fail to completely rescue females from the lethal effects of loss-of-function *Sxl* and *da* mutations which are considered to be due to inappropriate X chromosome hyperactivation (Skripsky and Lucchesi, 1982; Uenoyama et al., 1982; Cline, 1984).

(2) The lethal periods of these mutations are later than the lethal periods of *Sxl* and *da* (Linsley and Grell, 1968; Tanaka et al., 1976; Belote and Lucchesi, 1980; Uchida et al., 1981).

(3) An assay for dosage compensation based on the mutant phenotype of a hypomorphic allele of the X-linked *runt* locus revealed no effect of these male-specific lethal mutations on early dosage compensation (Gergen, 1987).

Two loci have been identified, on the evidence of their mutant phenotypes, as possible mediators of the dual functions of the *Sxl* gene in dosage compensation and sex determination. These are *virilizer* (*vir*) and *hermaphrodite* (*her*). The sex-transforming effects of certain *virilizer* alleles are seen only in chromosomal females,

which develop as intersexes; however, there are lethal alleles that affect both sexes. The female lethal allele *vir*^{H2} rescues both the lethal and sex-transforming effects of *Sxl*^{M#1} in XY flies, suggesting that *vir* acts downstream of *Sxl* in females to mediate both the female pathway of differentiation and a reduced rate of X chromosome transcription.

Support for a *vir* role in female dosage compensation came from the observation that the female lethal effect of *vir*^{H2} is partially rescued by male lethal mutations in genes that are required for hypertranscription of the single male X chromosome. The XX flies rescued by this interaction exhibit the masculinizing effects of the *vir*^{H2} mutation (H. Amrein et al., unpubl.). The existence of an allele that is lethal to both sexes suggests there may be an additional vital function for this gene that remains to be determined.

The *her* locus was identified as a temperature-sensitive recessive mutation that transforms females into intersexes, causes slight feminization of male flies, and reduces the viability of both sexes greatly. Both aspects of the mutant phenotype are influenced by the maternal genotype at this locus (G. Carson, pers. commun. in Baker and Belote, 1983). These observations suggest that the *her* product may be required for sex determination and dosage compensation in both sexes (Baker and Belote, 1983), although the current model of the regulatory hierarchy does not readily accommodate such a gene. In having an active role in male and female sex-determination, *her* is reminiscent of the *dsx* locus, and one could speculate that it is involved in either mediating or activating the dual *dsx* functions and in some nonsex-specific vital process. Further investigation of *her* is obviously required to clarify its function.

Sex-Lethal Regulation Is by Means of Sex-Specific RNA Splicing

The *Sxl* region has been cloned by the P-element transposon tagging method (Maine et al., 1986). Restriction fragments from this region were used to probe Northern blots and revealed a complex pattern of transcription from this locus. The gene contains several introns, and a variety of splicing patterns are used to generate multiple, overlapping transcripts, in-

cluding male-specific and female-specific size classes.

Three transient *Sxl* transcripts were detected at the blastoderm stage, later being replaced by a second set of six transcripts, which persists throughout the remainder of development (Bell et al., 1988). Sex specificity of the early transcripts has not been determined, but three of the later transcripts were shown to be female specific (1.9, 3.3 and 4.2 kb) and three to be male specific (2.0, 3.6 and 4.4 kb). Within each sex the three later transcripts show extensive overlap and share most exons, but differ in length at their 3' ends. The male transcripts appear identical to their female counterparts except for containing an additional third exon. Examination of adult male- and female-specific *Sxl* cDNAs confirmed and extended the results of the Northern blot analysis. Comparison with the genomic sequence revealed that both cDNAs extended over 15 kb of the 23 kb transcribed region and possessed identical 5' ends. The most significant difference between the male and female cDNAs was the presence in the former of an additional 190 bp third exon (making a total of eight to the female's seven) after base 599 of the female clone; this exon introduces two stop codons in the long open reading frame of the female transcript, such that the male transcript would encode only the first 48 amino acids of the 354-amino acid female product (Fig. 1).

Other minor differences between the two cDNAs in the splice sites of the first, second and fifth exons were shown by RNAase protection analysis to be nonsex-specific, and examination of other cDNA clones has confirmed this (M. Samuels, pers. commun. in Bell et al., 1988). Probing of Northern blots with a male third exon-specific probe confirmed that the extra third exon was present in all three adult male transcripts but none of the female transcripts, although RNAase protection studies with five different male cDNA clones revealed that some of them contained a smaller 172 bp version of this male-specific exon, and sequence data showed that the smaller exon arose from the use of a splice acceptor site 18 bp farther downstream. The same two termination codons are introduced by the smaller exon, however, resulting in an even shorter predicted protein product of 42 amino acids.

The male and female *Sxl* cDNA clones ex-

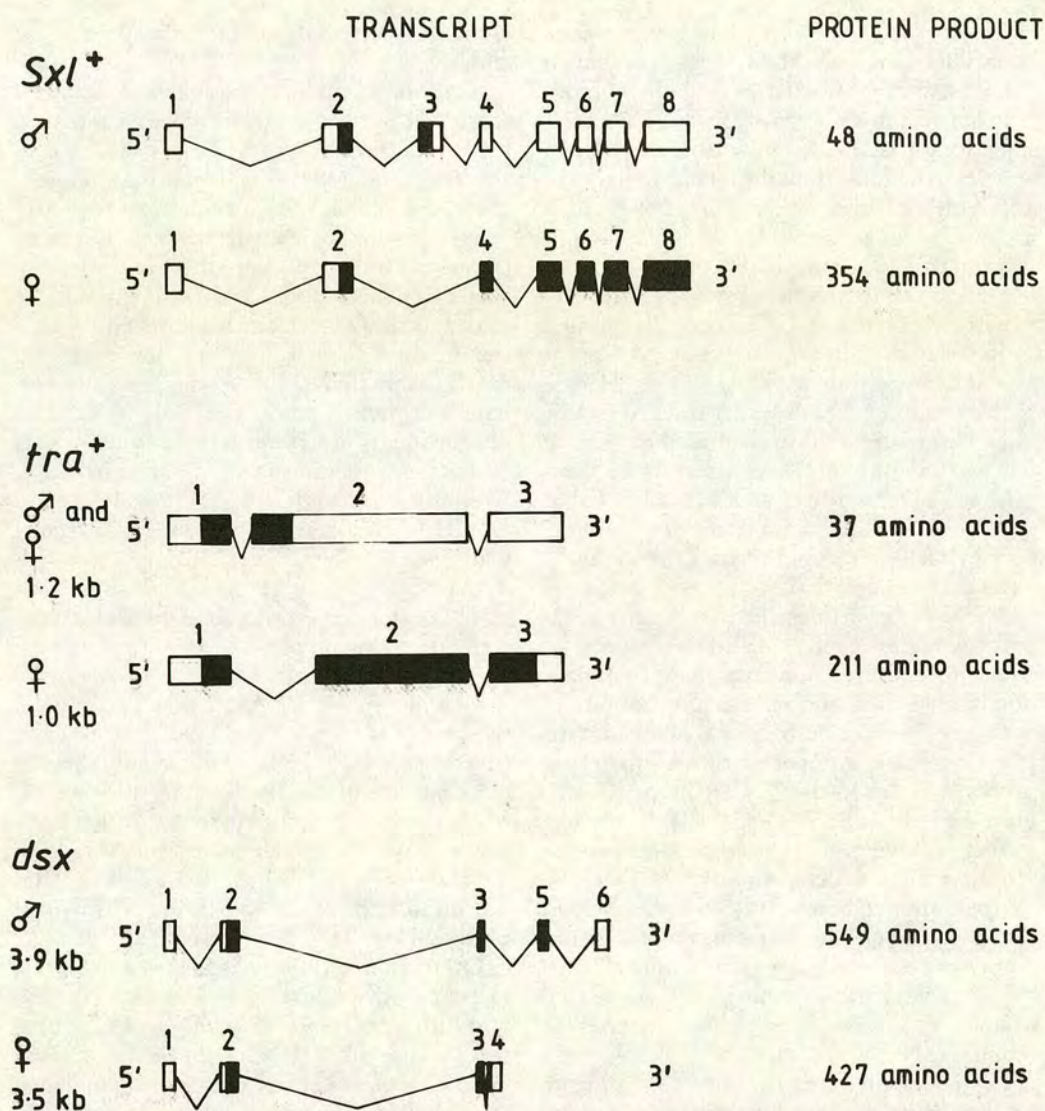


FIG. 1. SEX-SPECIFIC RNA PROCESSING IN THE REGULATORY HIERARCHY OF SOMATIC SEX DETERMINATION

Shaded areas represent coding sequence. In adult males and females, there are actually several size classes of *Sxl* transcript differing in length at their 3' ends (see text). The structures depicted correspond to sex-specific cDNAs which were truncated at their 3' ends and so could not be assigned to particular size classes (Bell et al., 1988).

In male flies *dsx* also produces a 2.9 kb transcript that probably has identical coding capacity to the 3.9 kb transcript (see text).

Female-specific splicing of the *Sxl* transcript results in an active protein that dictates the female-specific *tra* splicing pattern and the resulting *tra* protein dictates the female-specific *dsx* splicing pattern. Male-specific splicing of the *Sxl* transcript truncates the reading frame and no active protein is produced. In the absence of an active *Sxl* product, the *tra* transcript is spliced in the non sex-specific mode, the reading frame is truncated and no active *tra* protein produced. In the absence of an active *tra* protein the *dsx* transcript is spliced in the male specific mode. The male and female-specific *dsx* transcripts both encode proteins, with active and opposite roles in sex-determination

amined were both truncated, preventing precise determinations of the 3' ends of their corresponding transcripts; they could not therefore be assigned to a particular size class. It seems likely that they result from differential splicing of a common RNA precursor, however, and since the additional male exon can account for the observed size difference between all corresponding male and female *Sxl* transcripts, it probably constitutes the only sex-specific feature of the *Sxl* transcripts.

This molecular analysis of the *Sxl* locus suggests therefore that the gene's active state in females and inactive state in males, inferred from the genetic studies (see earlier) are achieved by sex-specific RNA splicing events, such that an active protein is produced only in the female. In view of the proposed inactivating function of the male-specific exon, it seems significant that all five constitutive gain-of-function male-lethal alleles of *Sxl* have transposons inserted within 1 kb of this exon (Maine et al., 1986), and it may be supposed that these insertions somehow disrupt the normal splicing pattern and promote female-specific splicing in chromosomal males (Bell et al., 1988).

As to the question of what determines the appropriate *Sxl* RNA splicing patterns in the two sexes, it seems likely that some product of the early *Sxl* transcripts has a role to play, since they are present only at the blastoderm stage, when genetic evidence suggests that the state of *Sxl* activity is set, becoming independent of any signal thereafter (Sanchez and Nöthiger, 1983; Cline, 1984; 1985). It could be envisaged that in female embryos some interaction, involving X chromosome numerator elements (such as *sis-a* and *b*) and maternal (*da*, *Dk*, *fs(1)1621*) and zygotic (*fs(1)1621*) gene products, results in the production of an early female-specific *Sxl* product which, in turn, promotes the adult female pattern of *Sxl* RNA processing. On the other hand, in male embryos the lower dosage of X chromosome numerator elements results in some alternative early *Sxl* transcription pattern and the male-specific later pattern.

The 5'-most exon of the adult *Sxl* transcripts was found to be absent from the early transcripts, suggesting that they are initiated at a different site (Salz et al., 1989); this possibility is supported by the fact that deficiencies that remove the 5'-most adult exon, and are defec-

tive in adult functions, complement a *Sxl* allele that lacks exclusively early functions (Maine et al., 1986; Cline, 1986). Also it appears that the male-specific exon is spliced out of these early transcripts. In the light of these findings it has been proposed that early *Sxl* regulation may be transcriptional rather than post-transcriptional, with early transcription confined to female embryos (Salz et al., 1989).

Whatever the precise mechanism determining the alternative splicing of the precursor of the later transcripts, some distinctive features of the alternative splice sites, reflecting the nature of the mechanism, might be predicted. In fact, however, all the regulated splice sites of the adult *Sxl* transcripts conform to the consensus sequences (Breathnach and Chambon, 1981; Bell et al., 1988) and provide no clues as to how the mechanism operates.

Although the proposed early female-specific *Sxl* product could direct excision of the male-specific exon from the adult transcripts, the question remains as to what governs its removal from the early transcripts. One possibility is that maternally derived *Sxl* transcripts, which are detected in early embryos, are responsible (Salz et al., 1989).

Since the *Sxl* gene is autoregulatory, maintaining its appropriate expression independently of the initiating signal (Sanchez and Nöthiger, 1983; Cline, 1984), it follows that the adult female-specific gene product must be able to control the female-specific RNA splicing pattern, and in this connection it is significant that the predicted protein product of the female *Sxl* cDNA clone shows considerable homology to several known ribonucleoproteins (Bell et al., 1988).

In addition to its autoregulatory function, *Sxl* is necessary for female sexual differentiation and dosage compensation (see above). Since *Sxl* appears to autoregulate by means of differential RNA splicing, it might be expected that it regulates the other genes under its control by the same mechanism; indeed, it has been shown recently that *tra*, the gene below *Sxl* in the somatic sex determination hierarchy (Baker and Ridge, 1980; McKeown et al., 1988), is also regulated at the level of differential RNA processing (Boggs et al., 1987; McKeown et al., 1988) which again involves the splicing out of translation termination codons to produce an active female-specific product. Intriguingly, a

long T tract exists in the vicinity of the regulated splice acceptor sites of both *Sxl* and *tra*, suggesting a model whereby the female *Sxl* product binds near the regulated splice sites of the *Sxl* and *tra* RNAs, blocking their utilization and promoting utilization of the female-specific splice acceptors (Bell et al., 1988); recent analysis of the control of sex-specific splicing of *tra* RNA has provided strong support for such a model of *Sxl* action (Sosnowski et al., 1989). At least one additional gene of the hierarchy of somatic sex determination, *dsx*, is also regulated at the level of RNA splicing by an upstream gene product or products. The regulation of *tra* and *dsx* is discussed in detail below.

Sxl is also required for female germ line differentiation (Schüpbach, 1985; Steinmann-Zwicky et al., 1989), and since *Sxl* mutations specifically defective in a germ line function have been identified (Salz et al., 1987), it seems significant that two of the adult female *Sxl* RNAs are apparently specific to the female germ line. One of these is the 1.9 kb transcript. In addition, in females the 3.3 kb transcript size class was found to consist of two distinct RNA species, one of which is germ line specific. The germ line specificity of these transcripts was determined by comparing their abundance in wild type females and mutants with defective or absent germ lines (Salz et al., 1989). It remains to be established whether these transcripts encode germ line-specific polypeptides, which would indicate a functional distinction between *Sxl* activity in germ line and soma; but since the 4.2 kb transcript and one of the 3.3 kb adult female transcripts seem to be present in soma and germ line, at least some aspect of *Sxl* function is probably common to both. While the role of *Sxl* in somatic sex determination is well characterized, the genes that *Sxl* might interact with in the germ line are unknown; and although several genes involved in male dosage compensation have been identified by their male-lethal mutant phenotypes and are under negative *Sxl* regulation (see earlier), this regulation has not been characterized at the molecular level. As information becomes available it will be interesting to see if sex-specific RNA splicing is a common mechanism underlying all *Sxl* functions.

THE AUTOSOMAL REGULATORY HIERARCHY

In addition to *Sxl*, four autosomal regulatory genes, *transformer* (*tra*), *transformer-2* (*tra-2*), *doublesex* (*dsx*) and *intersex* (*ix*), because of the sex-transforming effects of mutations at these loci, have been identified as mediators of the choice between male and female developmental pathways. Their relative positions in the linear hierarchy of sex determination have been established. Although the *her* and *vir* loci evidently have a role in sex determination as well, their characterization is at an early stage and their positions within the hierarchy are uncertain. For these reasons, they are not considered further.

Baker and Ridge (1980) used X ray-induced mitotic recombination to generate phenotypically marked cell clones that were simultaneously homozygous mutant for all possible combinations of the genes of the hierarchy and then observed the effects on sexual differentiation. From the patterns of epistasis observed, they concluded that these loci form a linear hierarchy as follows:

$$Sxl \rightarrow tra, tra-2 \rightarrow dsx$$

This means that *dsx* is epistatic to all other mutations, and that flies homozygous for a recessive *dsx* mutation and any other mutation have identical phenotypes. Subsequent studies have largely confirmed these conclusions and only the position of *ix* within the hierarchy remains to be established with certainty (Nöthiger et al., 1987). A working model of how these loci interact to control sexual differentiation, based upon the mutant phenotypes listed in Table 1 and the observed pattern of epistasis, is as follows (Baker and Belote, 1983): in chromosomal females the product of the activated *Sxl* gene acts to "turn on" the *tra* and *tra-2* loci, the products of which act together to induce the female mode of expression of the bifunctional *dsx* locus, resulting in the repression of male differentiation functions and permitting female functions. If the chromosomal sex is male, *Sxl*, *tra* and *tra-2* are inactive and *dsx* remains in its basal mode of expression, repressing female differentiation and permitting male differentiation. The *ix* locus appears to have no role in determining the mode of expression of *dsx* but

its product, in conjunction with the *dsx* female product, is needed in the female for repression of male functions.

The regulatory genes of the hierarchy governing sexual differentiation show several properties in common with the homeotic regulatory loci of *Drosophila* that control segment identity, in that they act in a cell-autonomous manner to mediate binary decisions between alternative developmental pathways (Lewis, 1978; Kaufman et al., 1980; Belote et al., 1986). Also, like the homeotic loci, the genes of the sex determination regulatory hierarchy are required at several stages in development, as discussed below.

In the following sections, genetic and recent molecular data on the autosomal genes of the hierarchy are presented in some detail, and the implications of the new molecular data for the operation of the regulatory hierarchy are discussed. A model of the hierarchy, which accommodates the new data, is shown in Fig. 2.

transformer and *transformer-2*

Null mutations at the *tra* (Brown and King, 1961) and *tra-2* (Watanabe, 1975; Belote and Baker, 1982) loci, when homozygous in chromosomal females, cause them to develop somatically as pseudomales, differing from normal males only with respect to their infertility because of incomplete transformation of their germ line. Neither mutation has any effect on somatic differentiation in chromosomal males, but male *tra-2* homozygotes do not form functional sperm, and temperature shifts with a temperature-sensitive *tra-2* allele suggest that the gene may function in the interval between late primary spermatocyte and early spermatid differentiation (Belote and Baker, 1983).

The genetic data point to *tra* and *tra-2* as being active in female somatic differentiation, translating the ON state of *Sxl* into expression of female-specific differentiation functions. Experiments involving the removal of wild type *tra* or *tra-2* function by X ray-induced mitotic recombination in cell clones of chromosomal females have shown that *tra* and *tra-2* are needed until the early pupal stage for normal, female-specific pigmentation of the external cuticle of the 5th and 6th dorsal abdominal segments,

whereas in the foreleg these genes have fulfilled their function by 24 to 48 hours before pupariation. The more extended requirement for *tra*, *tra-2* function in the former tissue may simply reflect the higher rate of cell proliferation in abdominal histoblasts during early metamorphosis relative to the foreleg imaginal disc, so that in both tissues *tra*, *tra-2* function is required until 3 to 5 division cycles prior to cessation of division (Baker and Ridge, 1980). The requirement for these genes into the third larval instar/pupal period is consistent with the time of action of homeotic loci, as demonstrated by analogous experiments (Gehring, 1976; Lewis, 1978).

Studies with temperature-sensitive alleles of *tra-2* (*tra-2^{ts}*) have extended the above findings to show that functional *tra-2* product is required for normal differentiation in the foreleg until 2 to 4 days after the need for transcription of the gene ceases (Belote and Baker, 1982). These experiments, involving temperature shifts of XX;*tra-2^{ts}* homozygotes between the restrictive (29° C) and permissive (16° C) temperatures at various stages throughout development, also demonstrated that *tra-2* does not function to determine sex at a specific time point, but is probably required continuously within a single cell lineage to mediate a series of steps in the process of sexual differentiation. This conclusion followed from the observation that functional *tra-2* product is required from just before to just after pupariation to determine female bristle number in the last transverse row of the basitarsus of the foreleg, whereas their morphology (thin female basitarsal bristles or male sex-combs) is determined considerably later, at least 24 hours after pupariation.

As a step towards the elucidation of its activity at the molecular level, the *tra* gene has been cloned in a 250 kb chromosome walk covering the region to which it had been assigned by deletion mapping (Belote et al., 1985; Butler et al., 1986). Screening of cDNA libraries with fragments spanning the walk revealed that the gene encodes a 1 kb female-specific transcript. The *tra* transcription unit was sub-localized to a 2 kb DNA fragment and the technique of P element-mediated germ line transformation was used to demonstrate that this fragment could provide wild type *tra* function.

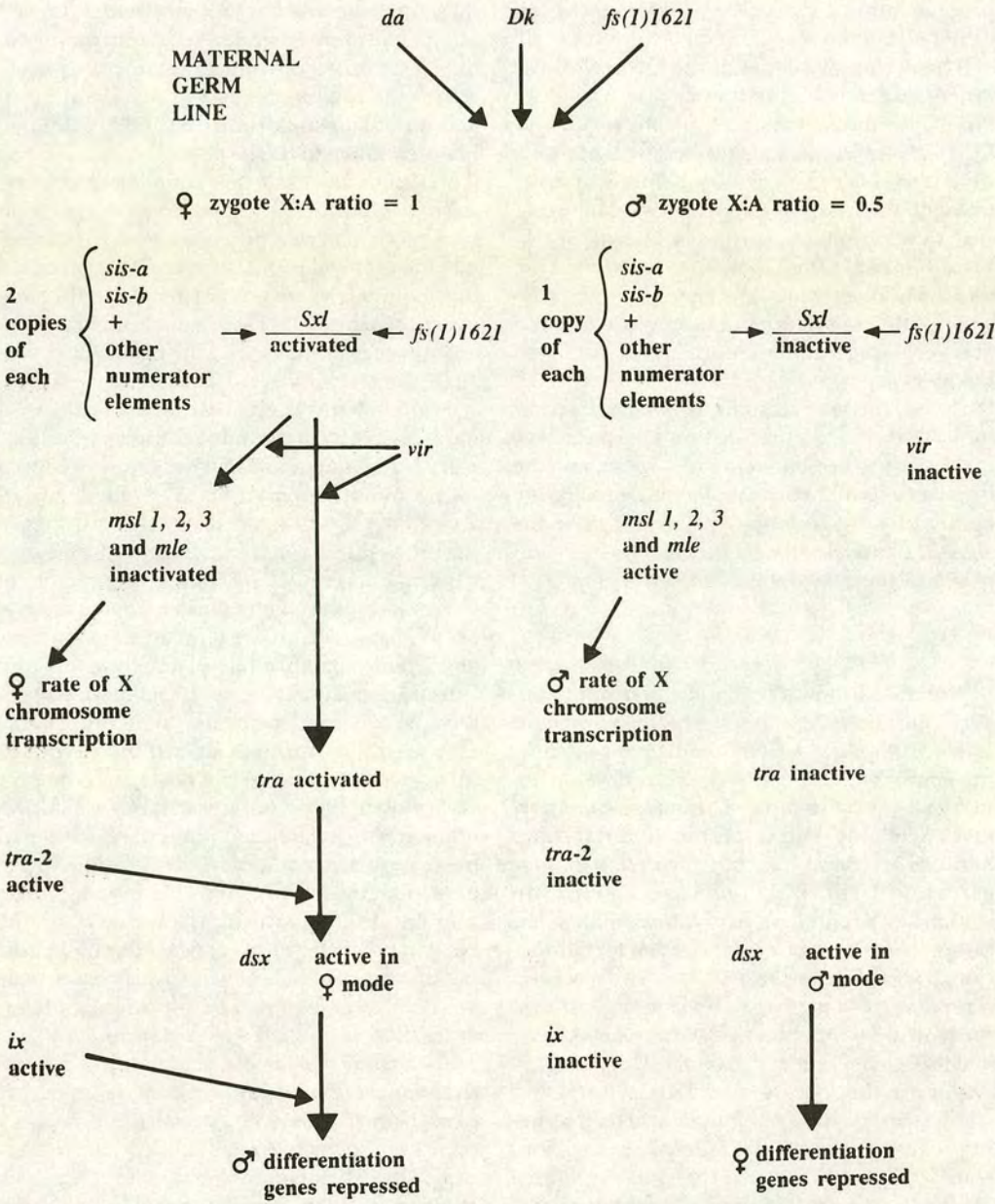


FIG. 2. THE REGULATORY HIERARCHY GOVERNING SOMATIC SEX DETERMINATION

This technique, which allows the introduction of defined DNA fragments into the *Drosophila* genome and observation of the resultant phenotype, is an invaluable research tool and provides the ultimate proof of a cloned gene's iden-

tity (Rubin and Spradling, 1982). In this case the 2 kb fragment demonstrated the capacity to fully rescue the female phenotype of XX;*tra/tra* flies, confirming that it contained the transformer gene and necessary regulatory

sequences (Butler et al., 1986; McKeown et al., 1987). Although rescue by the construct was shown to be only partial initially, this was shown to be a position effect rather than due to a lack of transcribed or *cis*-acting sequences, since transposition to a new location, induced by a dysgenic cross, yielded one line in which the *tra* phenotype was totally suppressed by a single copy of the 2.0 kb fragment. The small size of the *tra* locus is in contrast to the size of homeotic loci (e.g., *Ultrabithorax* is 70 kb) and other sex determination genes such as *Sxl* (23 kb — Bell et al., 1988) and *dsx* (> 40 kb — Baker and Wolfner, 1988).

More recent work has shown that the *tra* gene gives rise to a nonsex-specific 1.2 kb transcript in addition to the 1.0 kb female-specific transcript (Boggs et al., 1987). Using the cloned *tra* gene as a probe, cDNA copies of both size classes of transcript were isolated. Comparisons of the genomic and cDNA sequences revealed similar transcription initiation and polyadenylation sites for both mRNA size classes, implying that differing internal RNA splicing events account for the size difference. This was shown to be due to alternative splicing of the first intron, such that the nonsex-specific mRNA results from the use of a 3' splice acceptor 125 bp upstream of the female-specific acceptor site. Sequence data also revealed that the nonsex-specific message contains no long open reading frame, while the female transcript encodes a polypeptide of 211 amino acids (Fig. 1).

The same study determined that both *tra* transcripts are present at all stages of development from embryo to adult; this is surprising, since temperature shifts of *tra-2^{ts}* homozygotes at early developmental stages have no detectable effects (Belote and Baker, 1982), and no sex-specific *dsx* transcripts are detected at these stages (Baker and Wolfner, 1988). It may be significant, however, that maximum levels of *tra* mRNA were detected at the pupal stage, when the most overt sexual differentiation is occurring. P-element transformation studies have revealed that the sex-specific splicing of the common precursor *tra* mRNA is independent of any upstream regulatory sequences and does not involve, for example, alternative promoter usage. Specifically, when XX;*tra/tra* flies are transformed with a construct comprised of the *tra* coding region fused at -19 from the transcription initiation site to a *Drosophila* heat shock

protein 70 (hsp70) promoter, partial rescue of the female phenotype was observed. The incompleteness of rescue could be attributed to genomic position effects on the inserted construct (Boggs et al., 1987).

P-element constructs fusing the constitutively active hsp70 promoter to *tra* cDNAs have been useful in further elucidating the relationship between the two *tra* mRNA species and have confirmed the role of the female-specific gene product in sex determination. Initially, the possibility that the nonsex-specific mRNA was a precursor of the female-specific mRNA was tested by transforming XX;*tra* individuals with a construct fusing the hsp70 promoter to the nonsex-specific cDNA. Since no rescue of the *tra* phenotype was observed, it was reasoned that the nonsex-specific message is not a precursor, confirming that differential splicing of a common precursor was involved as outlined above. Using a construct in which the hsp70 promoter and the female-specific cDNA are fused, it is possible to obtain ectopic expression of the *tra* female-specific gene product in flies of XY constitution. XY flies transformed with this construct exhibited full female somatic differentiation. As expected, these flies were sterile due to nontransformation of their germ line. Thus, the female-specific *tra* mRNA provides sufficient *tra* function to induce female differentiation and the nonsex-specific *tra* mRNA is unnecessary for female development. Since male flies homozygous for the original *tra* deletion (Sturtevant, 1945) do not produce detectable levels of either *tra* transcript, but are viable and fertile, the nonsex-specific *tra* mRNA is also evidently without function in males.

The ability to obtain ectopic expression of the *tra* female-specific product in male flies, thus mimicking a dominant gain-of-function allele, has proven valuable in confirming the linear nature of the regulatory hierarchy governing sex determination. Since the female-specific *tra* product induces female differentiation in XY flies, it may be inferred that any other genes required for female differentiation either act between the X:A ratio and *tra* or are dependent on *tra*. This result is not consistent with alternative models envisaging a branched regulatory system, in which the sex-specific products of other genes regulated independently of *tra* are essential components.

Additional corroboration for the linear model and the gene order

$$Sxl \rightarrow tra \rightarrow tra-2 \rightarrow dsx \rightarrow ix$$

comes from experiments that show that the *tra* female product is capable of complementing mutations in putative upstream genes, but depends on functional copies of downstream genes. Standard genetic crosses were used to introduce the ectopic *tra* female-specific function into flies homozygous for a masculinizing loss-of-function *Sxl* allele, and all such flies developed as females, a result that places *Sxl* unequivocally upstream of *tra* in the hierarchy. Further evidence for the upstream position of *Sxl* comes from molecular studies showing that *Sxl* is necessary for the production of the female-specific *tra* mRNA (Nagoshi et al., 1988).

Another cross produced flies that lacked *tra-2* function but carried the *hsp70/tra* female cDNA construct. All such flies, whether of XX or XY constitution, developed male morphology in accordance with the *tra-2* mutant phenotype, which is consistent with the above gene order. These results, however, do not exclude the possibility that *tra-2* does not exhibit a sex-specific pattern of expression, and is thus expressed in both sexes but only active in the presence of the *tra* female-specific product. Recent molecular evidence supports the idea that *tra-2* is on a side branch of the hierarchy (see section on *tra-2*).

Genetic studies (Baker and Ridge, 1980) have suggested that *dsx* is below *tra* in the hierarchy, in which case constitutive expression of the *tra* female product should not overcome the effect of loss-of-function mutations at this locus. Appropriate crosses showed this to be the case. Both XX;- and XY;*dsx/dsx* flies carrying the *hsp70/tra* female cDNA construct retained the intersexual *dsx* phenotype. The caveat that applied to the conclusions regarding the position of *tra-2* in the hierarchy does not apply in the case of *dsx*, since genetic studies indicate that the gene is under sex-specific regulation and has different, active, functions in males and females; molecular studies have shown that the gene gives rise to sex-specific transcripts (Baker and Wolfner, 1988). Furthermore, the production of the female-specific *dsx* transcript has been shown to depend on functional *tra* and *tra-2* genes (Nagoshi et al., 1988). Perhaps the ultimate proof that *tra* governs the sex-specific

expression of *dsx* comes from the direct molecular evidence of an experiment in which the *hsp70/tra* female cDNA construct was shown to cause *dsx* to be transcribed in its female mode in XY flies (McKeown et al., 1988).

The *ix* gene is believed to act downstream of, or in conjunction with, the *dsx* female gene product and thus cause female differentiation (Baker and Ridge, 1980). Accordingly, the induction of female differentiation by the *hsp70/tra* female cDNA should depend on a functional *ix* gene product, and this was shown to be the case (McKeown et al., 1988). Confirmation that *ix* acts at the bottom of the hierarchy comes from the observation that in homozygous *ix* intersexes only the female-specific *dsx* transcript is produced (Nagoshi et al., 1988); but, as in the case of *tra-2*, *ix* could be on a side branch rather than being under the regulation of the hierarchy.

Genetic and molecular experiments have shown that the female-specific *Sxl* product acts directly or indirectly to control expression of the *tra* gene (Cline, 1979; Nagoshi et al., 1988; McKeown et al., 1988). It has been hypothesized, therefore, that in female flies the products of the *Sxl* gene or an intervening, as yet unidentified, gene interacts with the splicing machinery to generate the female-specific *tra* splicing pattern. This could be achieved in two ways: either by interference with splicing at the nonsex-specific site such that the female-specific site is used instead (negative control) or by acting positively to promote use of the female-specific splice acceptor. Results from P-element transformation experiments, using *tra* constructs designed to distinguish between these two possible mechanisms, suggest that the negative control model is correct. Initially, a fragment comprising the female-specific intron plus 5' and 3' splice sites was placed in the context of the herpes simplex virus thymidine kinase gene; flies transformed with this construct exhibited correct sex-specific regulation of 3' splice site choice, demonstrating that sequences mediating this regulation are contained within the female-specific intron.

A prediction of the negative control model of splice control is that elimination of the nonsex-specific splice acceptor should result in utilization of the female-specific splice acceptor in males (which lack *Sxl* activity), and this was shown to occur. Flies were transformed

with a deletion derivative of the *tra* gene, which lacked the sequence surrounding the nonsex-specific splice acceptor, from the putative branch point to beyond the AG splice site, and the construct was found to induce an equal degree of female differentiation in both XX and XY flies which were deleted for the endogenous *tra* gene. RNAase protection analysis confirmed that the female-specific splice was operating at the same level in flies of both chromosomal sexes, and was therefore independent of *Sxl* activity. Furthermore, as predicted, deletion of the female-specific splice site resulted in female-specific accumulation of unspliced *tra* RNA. These results are entirely consistent with a negative control mechanism and difficult to reconcile with a mechanism in which *Sxl* acts positively to promote use of the female-specific splice acceptor. As noted earlier, similar sequences have been found in the vicinity of the *Sxl* male-specific 3' splice site and the *tra* nonsex-specific 3' splice site, which could represent a recognition element for a negatively acting regulatory protein. To test this possibility Sosnowski et al. (1989) mutagenized this putative recognition sequence in the *tra* gene and looked at the effect on splicing. The mutagenized construct showed no capacity to induce female differentiation in XY flies or XX flies deleted for *tra*; RNAase analysis revealed only a very low level of the female-specific *tra* RNA in the transformed lines, and a relatively high level of the nonsex-specific RNA. Together, these transformation experiments provide strong evidence that sex-splicing of *tra* RNA is under negative control, involving sequence specific interaction of a protein with the region of the nonsex-specific splice acceptor to partially block its utilization in females. The *Sxl* product is an obvious candidate for the proposed regulatory protein, perhaps complexed with a nuclear RNA in the form of a snRNP, but other factors could also be involved: for example, the product of the *virilizer* gene or even an as yet unidentified gene acting downstream of *Sxl* (Sosnowski et al., 1989).

The *tra-2* gene has now also been cloned and sequenced. Northern blots indicated a single transcript of 1.7 kb present at equal levels in both sexes. Isolation and sequencing of cDNAs indicated a *tra-2* protein product of 179 amino acids with homology to several RNA binding and splicing proteins (Amrein et al., 1988;

Goralski et al., 1989). It was recently reported, however, that *tra-2* actually gives rise to three distinct 1.7 kb messages which result from differential RNA splicing and encode proteins with identical carboxy termini, including a putative RNA binding domain, and variable amino termini (W. Mattox et al., unpubl.). Two of the transcripts are nonsex-specific and the third is male specific. The absence of a female-specific *tra-2* transcript argues against regulation of the gene by *tra*, at least at the level of transcription or RNA processing, since *tra* is only active in female flies (see earlier) and supports a model in which *tra-2* is on a side branch of the hierarchy (Fig. 2).

A plausible explanation to accommodate these data would be that the *tra-2* product involved in female differentiation is present in both sexes but only active in conjunction with another female-specific gene product, perhaps that of *tra* (Fig. 2), whereas the male-specific transcript encodes a function involved in male germ line development, consistent with the sterility of XY;*tra-2/tra-2* flies. P-element transformation of *tra-2* mutant flies with the various *tra-2* cDNAs is being employed to determine the function of each transcript, and should reveal whether this model is correct (T. J. Goralski et al., unpubl.).

If, as the genetic and molecular evidence suggests, the products of *tra* and *tra-2* act together to cause female-specific splicing of a *dsx* primary transcript, it seems significant that the common domain of the predicted *tra-2* proteins appears very similar to other proteins known to be involved in RNA binding or splicing. In addition, P-element transformation with a construct fusing the 5' portion of the *tra-2* gene to part of the *E. coli lacZ* gene, followed by histological staining, showed β -galactosidase activity localized to the nucleus in several tissues, consistent with the proposed action of *tra-2* product in RNA splicing (W. Mattox et al., unpubl.).

doublesex

The *dsx* locus is believed to be differentially active in males and females, mediating either male or female terminal differentiation functions (Table 1) (Baker and Ridge, 1980). This is inferred from the phenotypes of various mutant alleles at this locus. Recessive *dsx* mutations transform either chromosomal males,

chromosomal females, or both into intersexes exhibiting intermediate sexual differentiation at the level of individual cells, which seem to be expressing male and female functions simultaneously (Hildreth, 1965; for reviews see: Baker and Belote, 1983; Nöthiger and Steinmann-Zwicky, 1985; Nöthiger and Steinmann-Zwicky, 1987). This is in contrast to the mosaic pattern of male and female tissues characterizing the 2X:3A intersexes referred to earlier, in which cells choose either the male or female pathway.

The fact that null *dsx* mutations affect both males and females suggests that this gene is active in both sexes, whereas the existence of sex-specific *dsx* alleles indicates that the gene functions differently in the two sexes. Molecular analysis has proven consistent with this interpretation of *dsx* function (see below). A model of *dsx* function, consistent with the genetic and molecular data on *tra*, *tra-2* and *dsx* mutations, envisages that the ground state of activity of the gene (i.e., in the absence of *tra* or *tra-2* product) is that found in males, when the gene product acts to repress female differentiation functions. In the female the product of the *tra* and *tra-2* genes are continuously required to prevent expression of the *dsx* gene in the male mode and permit expression of the female product, which represses male differentiation functions (Baker and Ridge, 1980; Belote and Baker, 1983; Nöthiger et al., 1987; Nagoshi et al., 1988).

The existence of two dominant *dsx* alleles, *dsx^D* and *dsx^{Mas}*, which transform XX flies into intersexes, facilitated the cloning of the *dsx* region (Belote et al., 1985). After X ray mutagenesis of flies carrying these alleles, the F₁ generation was screened for revertants of the dominant phenotype. Of 32 such revertants obtained, 13 on cytological examination were found to have inversions or translocations broken in salivary gland chromosome band 84E1-2, placing *dsx* in this region. One deficiency moved *dsx* near a previously cloned α -tubulin gene, which facilitated a "jump" into the *dsx* region; a 107 kb chromosome walk was then conducted, covering the extent of the *dsx* region necessary for its male-specific function, as defined by deficiency and translocation breakpoints.

Probing of Northern blots with clones from this region identified multiple, overlapping

transcripts and a complex pattern of temporal and sex-specific regulation (Baker and Wolfner, 1988). Two nonsex-specific transcripts of 1.65 kb and 2.8 kb are present throughout the larval period. At the end of the larval period, these transcripts are replaced by male-specific transcripts of 2.9 and 3.9 kb and a female-specific transcript of 3.5 kb (Baker and Wolfner, 1988; Nagoshi et al., 1988; Burtis and Baker, 1989). In adults, low levels of the sex-specific transcripts persist, along with a 0.7 kb male-specific transcript (Baker and Wolfner, 1988).

It appears that the larval nonsex-specific transcripts result from differential splicing of a primary transcript, since they are both homologous to the same genomic fragments at their 3' and 5' ends but show different patterns of homology to internal fragments.

In the case of the male- and female-specific transcripts, Northern blot analysis again suggested that they were generated by differential splicing of a common precursor RNA; this was confirmed by cloning and sequencing of cDNAs corresponding to the 3.9 kb male-specific and 3.5 kb female-specific transcripts and comparison with genomic sequence (Burtis and Baker, 1989). The transcripts share their first three exons, of 709, 1367 and 138 nt. The female-specific transcript possesses a single additional 3' exon of 1168 nt, following a small intron of 114 nt, while the male-specific transcript possesses two male-specific exons following introns of 4.1 kb and 4.9 kb respectively. In addition, the male- and female-specific transcripts utilize different polyadenylation sites (Fig. 2). The minimum size of the *dsx* locus, defined by the extent of the male transcript, is greater than 40 kb. Although no cDNA which could be assigned to the 2.9 kb male-specific transcript was isolated, a canonical polyadenylation signal was found close to the position of the last male exon expected for a message of that length.

All the male- and female-specific *dsx* messages contain single long open reading frames, beginning in the second common exon and encoding polypeptides of 549 amino acids (57.4 kd) and 427 amino acids (44.8 kd) respectively. Both proteins share their first 397 amino acids but there are male- and female-specific carboxy-terminal domains of 152 and 30 amino acids respectively. The amino acid sequences

yielded little clue as to the protein's mode of function and showed no significant homology to known DNA binding motifs, although they do contain strings of repeated amino acids, as seen in other *Drosophila* regulatory proteins (Burtis and Baker, 1989).

The existence of sex-specific *dsx* products is consistent with the proposed mode of action of *dsx*, although the low abundance or absence of sex-specific transcripts in the larval stage seems to conflict with the finding that *tra-2* function is required as early as the second larval instar for normal differentiation of the genital primordium in females (Belote and Baker, 1982). As noted by Baker and Wolfner (1988), however, it is possible that female-specific *dsx* transcripts are being produced at earlier stages in only a few cells, and are therefore below the detection threshold in Northern analysis. Similarly, the low-level of sex-specific transcripts in adults may reflect continued low-level activity of the gene in certain tissues, such as in the female fat body, where yolk protein gene transcription is known to be dependent on the genes of the hierarchy (Belote et al., 1985;). In common with *Sxl* (Maine et al., 1986; Bell et al., 1988) and *tra* (Boggs et al., 1987; McKeown et al., 1988), *dsx* produces nonsex-specific transcripts that are not predicted by the genetic data. There is evidence that the nonsex-specific *dsx* transcripts may be without function, as is believed to be the case for *Sxl* and *tra* (see earlier). In the case of *dsx*, inversions and translocations which interrupt the region homologous to both larval and pupal transcripts have phenotypic effects that are indistinguishable from those of rearrangements which are broken in the portion of *dsx* that is homologous to only the sex-specific pupal transcripts. This implies again that these nonsex-specific transcripts are nonfunctional and incidental to the operation of the transcriptional machinery (Baker and Wolfner, 1988).

P-element transformation experiments, using *dsx* cDNAs under the control of an actin promoter, have confirmed that they encode the active products, since ectopic expression of either the male- or female-specific *dsx* products in flies of opposite chromosomal sex has sex-transforming effects (Burtis and Baker, 1989). In females, the male construct resulted in male-like pigmentation of the fifth and sixth abdominal tergites. Males were not visibly affected by

the female construct, but displayed substantially elevated levels of yolk protein gene transcripts, which are present only at very low levels in wild type males. The ability of these constructs to induce sex transformation and the existence of dominant masculinizing *dsx* mutations, needs to be reconciled with the genetic evidence that *dsx* acts as a repressor of the inappropriate differentiation pathway. It has been proposed therefore that the *dsx* products could normally function as homo-oligomers and that simultaneous expression of the male and female *dsx* products gives rise to nonfunctional hetero-oligomers and consequent failure to repress either male- or female-specific differentiation genes (Nöthiger et al., 1987; Burtis and Baker, 1989). Alternatively, both male- and female-specific *dsx* products may bind to *cis*-acting sequences associated with both male- and female-specific differentiation genes, but only repress the inappropriate genes. Accordingly, the ectopic expression in the above experiments would block access to the endogenous *dsx* products and permit inappropriate gene expression. Whatever explanation is correct, it must also account for the intersexual phenotype of XX flies heterozygous for the dominant masculinizing allele *dsx^{Mas}*, which has been shown to constitutively produce the male-specific transcript (Nagoshi and Burtis, unpubl.) and the wild type gene, for these flies also express both the male- and female-specific *dsx* functions (Burtis and Baker, 1989).

The same constructs, when introduced into flies that were homozygous for a loss-of-function allele of *dsx*, were unable to fully rescue the intersexual phenotype, but did produce a degree of sexual transformation in the predicted directions. The relative weakness of the phenotypic effects in all these transformation experiments was attributed to incorrect regulation by the actin promoter and other possible deficiencies of the constructs.

It seems from the evidence that everything that is necessary for male-specific splicing of *dsx* precursor RNA is present in females, but that the regulatory gene hierarchy is active in the female sex to impose the alternative splicing pattern. If this view is correct, then the expression of the *dsx* male function can be regarded as the default condition, independent of the other genes of the regulatory hierarchy.

Since there is no evidence to date that other

genes act between *tra*, *tra-2* and *dsx*, and in view of the fact that the predicted protein products of the *tra-2* gene has RNA binding potential (see earlier), it is tempting to speculate that the *tra* and *tra-2* gene products act together to impose the female-specific *dsx* splicing pattern. However, since the sex-specific *Sxl* (Maine et al., 1986) and *tra* transcripts (McKeown et al., 1987) are present long before the end of the larval period, when the sex-specific *dsx* transcripts first appear, it follows that there must be an additional temporal signal in both sexes that is required for sex-specific splicing to occur. A proposal consistent with the data is that at the end of the larval period a signal common to both sexes results in the production of a *dsx* precursor RNA, which is then spliced in a sex-specific manner directed by the *tra* and *tra-2* gene products (Baker and Wolfner, 1988).

Of all the splice acceptor sites used to generate the *dsx* RNAs, the female-specific splice acceptor site, within the 114 nt intron, shows the least similarity to the *Drosophila* consensus sequence, so the *tra* and *tra-2* products may act to promote the use of this relatively low-affinity site by the splicing machinery. In addition, the first half of the female-specific exon contains six repeats of a 13 nt sequence in a position to which all four dominant masculinizing *dsx* mutations have been mapped, which suggests that these sequences may have a regulatory function. Since the male and female *dsx* transcripts utilize different polyadenylation sites, however, it is also possible that choice of polyadenylation site is the primary regulatory event leading to differential splicing (Burtis and Baker, 1989).

In order to identify the sequences implicated in the control of sex-specific splicing at *dsx*, various sized DNA fragments, including the region of the sex-specific splice, have been fused to an heterologous promoter, introduced into flies by P-element transformation, and tested for splicing of the resultant transcripts (K.C. Burtis et al., unpubl.).

intersex

Mutations at the *ix* locus are recessive and, when homozygous, cause XX flies to develop as true (non-mosaic) intersexes, but have no visible effect on chromosomal males (Morgan et al., 1943). As previously discussed, molecular data have supported the genetic evidence

that *ix* acts below the *dsx* gene in the female differentiation hierarchy, but the possibility remains that *ix* is on a side branch (as *tra-2* appears to be) under separate regulation. In fact, there is some evidence that *ix* is also active in males. Although chromosomally male flies that are homozygous for an *ix* loss-of-function allele are visibly unaffected and fully fertile, they have been observed to elicit significantly more courtship behavior from other males than their *ix* siblings, suggesting that *ix* may have some role in regulating pheromone production in males (Tompkins, 1986); attempts to reproduce this male phenotype, using three *ix* alleles in homozygous, hemizygous and heteroallelic combinations, have proven unsuccessful (B. Chase and B. Baker, unpubl.), so the role of this locus in males remains open to question.

END OF THE LINE — THE SEXUAL DIFFERENTIATION GENES

The third level in *Drosophila* sex determination, after the assessment of the X:A ratio and its translation via the regulatory hierarchy into sex-specific expression of the *dsx* locus, consists of the structural, terminal differentiation genes, whose sex-differential expression is directly responsible for sexual dimorphism, and about which little is known. One would expect to find common *cis*-acting regulatory elements involved in the sex-specific expression of these genes, as predicted by Britten and Davidson's theory of transcriptional control in higher eukaryotes (Britten and Davidson, 1969) and since borne out by observation. Proceeding from the assumption that a multicellular organism composed of many specialized tissues would require that different subsets of the gene complement be active in different tissues, developmental stages and sexes, Britten and Davidson proposed a general model for gene regulation whereby structural genes are turned ON or OFF at appropriate times and places by the interaction of *trans*-acting factors with *cis*-acting elements associated with the structural genes, in a sequence-specific manner. This interaction would then induce or repress RNA polymerase II transcription of the gene.

Genes with similar patterns of expression are conceived of as belonging to batteries. Obviously there will be some overlap between batteries, with some genes belonging to more than one. In order to accommodate this require-

ment, two alternative models have been proposed. The first envisages redundancy at the level of the *cis*-acting elements flanking the structural genes, such that each structural gene possesses a *cis*-acting element for every gene battery to which it belongs. Expression of genes under complex regulation then requires the appropriate combination of *trans*-acting factors bound at their cognate *cis*-acting sites.

The existence of multiple *cis*-acting elements that flank genes would obviously require bound factors to exert their effect over a considerable distance. It has recently been hypothesized that the interaction of such remotely bound factors with the transcription complex could be facilitated by looping out of the intervening DNA, a phenomenon that has been successfully demonstrated *in vitro* with artificially separated lambda repressor binding sites (Ptashne, 1986; Schleif, 1987).

The second model envisages redundancy at the level of the *trans*-acting genes, such that a *trans*-acting gene exists for every gene battery, members of which have a single *cis*-acting site specific for that factor. The *trans*-acting factor is then required to have a pattern of expression identical to the genes it controls and the number of *trans*-acting genes is equal to the number of gene batteries with unique patterns of expression.

Studies of *cis*-acting elements that flank several structural genes suggest that they conform to the first model, possessing more than one *cis*-acting element, each mediating a different aspect of regulation (Gehring et al., 1984; Pirrotta et al., 1985; Garabedian et al., 1985; Shirras and Bownes, 1987).

Female-Specific Genes

The *cis*-acting sequences governing the expression of two sets of terminal differentiation genes expressed exclusively in adult females have been examined. These are the chorion protein genes and the yolk protein genes.

The multi-layered chorion of the egg, formed late in oogenesis, is the product of the sequential activation of batteries of genes in the ovarian follicle cells, where protein products are secreted. The huge demand for chorion proteins is met by the selective amplification of the genes (Spradling and Mahowald, 1980).

The X-linked S36 and S38 chorion protein gene transcription begins at stage 11 of oogen-

esis, peaking at stage 12, whereas the third chromosome cluster of S13, 15, 16, 18 and 19 genes is transcribed at stages 13 and 14. Individual genes within each cluster have unique mRNA accumulation profiles, however, rather than being under coordinate control. P-element transformation with constructs fusing various amounts of chorion gene upstream DNA to an alcohol dehydrogenase (*Adh*) reporter gene, determined that for the genes tested, all the *cis*-acting sequence necessary for correct stage tissue- and sex-specificity were contained within the first 1 to 2 kb of 5'-flanking DNA (although there was evidence that sequences either within the coding sequence or downstream of the S5 gene were required for transcription at the level of the endogenous gene). Thus an assay for *Adh* activity in the *Adh*-minus transformed flies indicated that reporter gene expression was precisely determined by the heterologous upstream DNA (Romano et al., 1988). In accordance with the model for redundancy at the level of *cis*-acting sequences, further analysis of the S15 gene 5' region has identified three adjacent regulatory elements, two of which mediate temporal specificity of expression, while the third element appears to determine tissue specificity (Wong et al., 1985; Mariani et al., 1988). For reasons that are discussed below, the chorion protein genes and other genes expressed exclusively in male or female gonadal tissue are probably not directly regulated by the somatic sex determination hierarchy, and are therefore unlikely to be associated with a binding site for the *dsx* gene product. The female-specific yolk protein genes, however, are regulated directly and one *cis*-acting region responsible for this regulation has been identified.

The three yolk proteins of *Drosophila melanogaster* are encoded by single-copy genes (YP1, YP2, YP3) on the X chromosome (Barnett et al., 1980). They are expressed in the fat body and ovarian follicle cells of adult females, and have provided a good system in which to study sex-specific, tissue-specific, developmentally regulated gene expression. YP1 and YP2 are separated by 1.2 kb and are divergently transcribed. YP3 is situated more than 1000 kb away (Hung et al., 1982). *Adh*-minus flies transformed with a YP1-*Adh* fusion gene, including 890 bp of YP1 5'-flanking DNA, exhibit a stage-, sex-, and fat body-specific pattern of

Adh activity characteristic of yolk protein genes. Furthermore, the construct was shown to be under the regulation of the somatic sex determination hierarchy, since no Adh activity was detected in chromosomal females transformed into pseudomales by a dominant *dsx* mutation (*dsx^D*) (Shirras and Bownes, 1987). Transformation studies with a marked YP1 gene, and including variable amounts of 5'-flanking DNA, have localized the sequence necessary for determining this expression pattern between -196 and -321 from the mRNA cap site. When this 125 bp region is fused upstream of an hsp70 promoter/*lacZ* construct, it confers heat shock-independent β -galactosidase expression of the same pattern, and it retains this effect when inverted or placed up to 8 kb downstream from the cap site (Garabedian et al., 1986).

Recently DNAase I footprinting studies, using extracts from embryos and a *Drosophila* Kc cell line, have determined that three different proteins appear to bind within a sub-region (-218 to -321) of the 125 bp sequence (P. Wensink, unpubl.). Which, if any, of these proteins is the male-specific *dsx* product has yet to be determined.

A *Drosophila* Kc cell line has previously been found to exhibit the female-specific splicing pattern of *dsx* RNA (Burtis and Baker, 1989). The implication from these footprinting studies, that the female-specific *dsx* product may bind within this YP gene regulatory element, has obvious relevance to the dominant sex-transforming effects seen in certain *dsx* mutations and ectopic *dsx* expression experiments (see earlier).

A 22 bp sequence from within this 125 bp region shows strong homology to a sequence downstream of the co-regulated YP3 gene (Garabedian et al., 1987), and part of this sequence also shows homology to a sequence 5' of the Larval Serum Protein 2 (LSP2) gene, which is expressed in the fat body of larvae and adults. It seems probable that these homologies reflect some co-regulation of the four genes, since a mutation (*cricket*) has been identified that causes reduced levels of all three YPs and LSP2 in adult flies (Shirras and Bownes, 1989).

Injection of flies with the steroid hormone 20-hydroxyecdysone induces YP gene expression in male flies (Postlethwait et al., 1980; Bownes et al., 1983) and increases transcrip-

tion in female flies (Jowett and Postlethwait, 1980; Bownes, 1986). Transformation experiments have also determined that the DNA sequences mediating this response lie outside the 890 bp 5' of YP1 (Shirras and Bownes, 1987), and in fact probably lie beyond the whole of the YP1 and YP2 structural genes with their intergenic spacer (Shirras, Bownes and Wensink, pers. commun.).

Analysis of the regulation of the YP genes is further complicated by the observation that their transcription is subject to nutritional modulation (Bownes and Blair, 1986), perhaps analogous to the blood meal-induced vitellogenesis observed, for example, in the mosquito, *Aedes aegypti* (Van Handel and Lea, 1984). P-element transformations of YP1-Adh fusion constructs have located the DNA required for this additional level of control within 890 bp 5' of YP1 (Bownes et al., 1988).

Ovarian follicle cell expression of the YP1 gene is mediated by *cis*-acting sequences separate from the 125 bp region discussed above. Whereas none of the YP1 P-element constructs discussed previously were transcribed in the ovaries, a YP2 construct including 342 bp of upstream DNA was transcribed in ovaries but not fat bodies (Garabedian et al., 1985). Further germ line transformation experiments have identified DNA between -159 and -340 of the YP2 cap site as being necessary for ovarian expression of YP1 and YP2. This region is sufficient to direct an ovarian pattern of expression from a heterologous promoter, but in an orientation-dependent manner, an intriguing finding since it must act bidirectionally on YP1 and YP2 (P. Wensink, unpubl.).

A second region influencing ovarian expression of YP1 was located within the first exon of YP2. It increases the level of YP1 transcripts in ovaries, but when fused to a heterologous reporter gene does not confer ovary- or stage-specific expression. This second region therefore behaves more like a general transcription enhancer. An indication of further complexity in YP gene regulation has come from the isolation of a DNA-binding protein with very high affinity for a 31 bp sequence 148 bp downstream of the transcription initiation site, which is necessary for normal YP1 RNA levels in vivo (P. Mitsis and P. Wensink, unpubl.). The regulatory regions defined by these transformation studies are illustrated in Fig. 3.

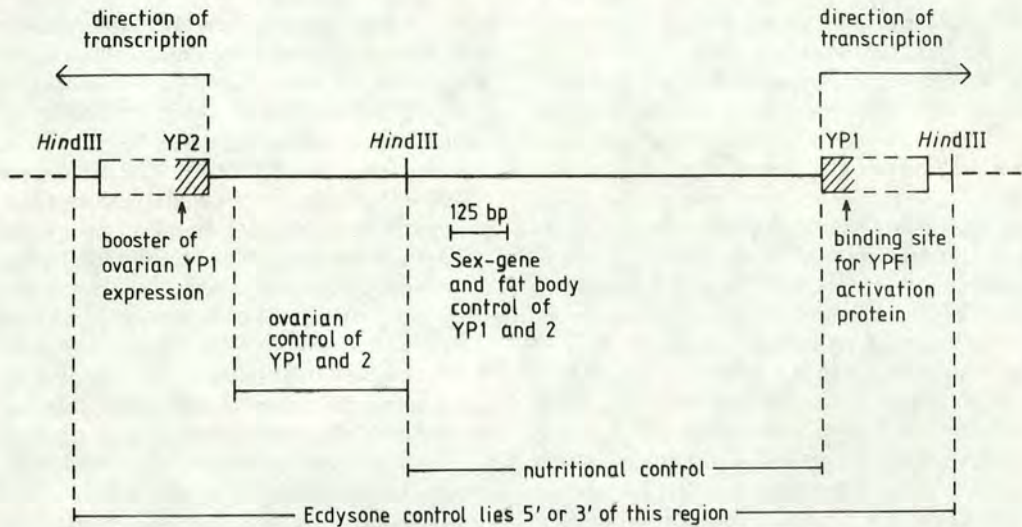


FIG. 3. DIAGRAM OF THE REGULATORY ELEMENTS GOVERNING EXPRESSION OF YOLK PROTEINS 1 AND 2 (YP1 AND YP2)

The observation that ovarian expression of YP2 is apparently independent of the 125 bp sequence shown to confer sex-, stage-, and tissue-specific expression on the adjacent YP1 gene, suggests that the hierarchy of regulatory genes determines YP gene expression by means of different mechanisms in the two tissues. Further evidence for this comes from the finding that in XX;*tra-2^{ts}* homozygotes, a shift from the permissive to restrictive temperature 1 to 2 days before eclosion abolishes fat body but not ovarian synthesis of yolk proteins, permitting some vitellogenesis to proceed (M. Bownes et al., unpubl.). Thus it appears that the synthesis of fat body yolk protein is dependent upon a functional *tra-2* gene product, while ovarian synthesis is not.

Confirmation that fat body YP gene transcription is continuously dependent on the regulatory hierarchy comes from experiments with fat bodies isolated from temperature-shifted XX;*tra-2^{ts}* homozygotes (Belote et al., 1985; Bownes et al., 1987). These experiments were, however, inconclusive with respect to dependence of the synthesis of ovarian yolk protein on the regulatory hierarchy.

The suggestion that ovarian synthesis does not depend on the continued function of the hierarchy, whereas fat body synthesis does, raises the possibility that sex-specific gene ex-

pression in sex-specific tissues may be irreversibly determined prior to adulthood and not continuously dependent on the hierarchy, whereas sex-specific gene expression in nonsex-specific tissues remains dependent on the hierarchy and thus sensitive to temperature shifts in *tra-2^{ts}* homozygotes. One could envisage that the process of differentiation leading to the formation of sex-specific tissues is initiated by the expression of *dsx* in the appropriate mode, while the expression of terminal differentiation genes in the formed tissue is mediated by a tissue-specific enhancer sequence (Wolfner, 1988).

Male-Specific Genes

By means of differential screening, several genes expressed in the male accessory gland have been isolated, and transcription from these genes has been shown to be insensitive to downward temperature shifts in adult *tra-2^{ts}* homozygotes (Di Benedetto et al., 1987; Monsma and Wolfner, 1988). Moreover, although the transcription of these three genes is not initiated until the late pupal stage, temperature shifts have shown that their expression is irreversibly determined during the late third larval instar. Bidirectional shifts have revealed that functional *tra-2* product need only be present during the late third larval instar

to determine later expression of these genes (Chapman and Wolfner, 1988).

Morphological examination of flies subjected to temperature shifts indicated a close correlation between the temperature-sensitive periods determining accessory gland formation and subsequent transcription of the three genes. Thus it seems reasonable that the male-specific expression of these genes is a consequence of an earlier decision to form accessory glands, dependent on a functional *tra-2* product (Chapman and Wolfner, 1988).

A number of genes expressed exclusively in the male germ line have also been isolated in differential screens and found not to be expressed in XX; *tra* or *tra-2* pseudomales, a finding that is consistent with the infertility of these flies being due to nontransformation of their germ line, as discussed previously (Di Benedetto et al., 1987; Schäfer, 1986).

A SEPARATE REGULATORY CASCADE FOR THE GERM LINE

While the regulatory hierarchy examined above appears to govern all aspects of somatic sexual differentiation, the same is not true of germ line differentiation. Chromosomally female *tra/tra* or *tra-2/tra-2* pseudomales have testes that contain abortive germ cells and are sterile (Seidel, 1963; Nöthiger et al., 1989). However, the construction of mosaics in which chromosomally female (XX) germ cells that were homozygous for the *tra* mutation were surrounded by female somatic tissue that was wild type at this locus (by means of pole cell transplantation) showed that the mutant cells could produce oocytes (Marsh and Wieschaus, 1978). Additional experiments showed that transplanted pole cells always developed according to their chromosomal sex, regardless of mutations at any of the autosomal regulatory loci, but only when the host had somatic cells of matching sex (Van Deusen, 1976; Marsh and Wieschaus, 1978; Schüpbach, 1982).

For spermatogenesis there is a germ-cell autonomous requirement for the Y chromosome, since XO males are sterile but XY pole cells transplanted into an XO male produce functional sperm (Marsh and Wieschaus, 1978). In the female the germ cells require a functional *Sxl* gene, as evidenced by the ovarian tumor phenotype of female germ line clones homozygous for a loss of function *Sxl* mutation (Schüp-

bach, 1985). It appears therefore that germ line differentiation involves cell-autonomous processes that are independent of the autosomal genes of the hierarchy of somatic sex differentiation, and in the male include Y chromosome-encoded functions. The failure of transplanted pole cells to develop in a somatic environment of opposite sex could simply reflect additional nongerm-cell-autonomous requirements for gametogenesis, contributed by the somatic environment, such as the yolk proteins in the case of oogenesis.

This picture of absolute germ line and somatic requirements is complicated somewhat by the fact that spermatocytes, spermatids and even immotile sperm have been observed in the testes of pseudomales, along with abnormal oocytes (Seidel, 1963; Nöthiger et al., 1989). Thus it appears that sexual pathway choice in germ cells may not be fully autonomous, and that the somatic cells can in some cases cause germ cells to differentiate according to their chromosomal sex in a somatic environment of phenotypically opposite sex.

Recent experiments involving pole cell transplantation have helped to clarify the situation (Steinmann-Zwicky et al., 1989). Whereas the earlier transplantation experiments showed that the generation of functional gametes required matching somatic and germ line sex, the fate of transplanted pole cells in a somatic environment of opposite sex was not determined because the donor cells were indistinguishable from host germ cells. Now, by using host flies that lack a germ line as a result of a maternal *oskar* mutation (Lehmann and Nüsslein-Vollhard, 1986), it has been shown that XX pole cells do enter spermatogenesis in a testis unless they carry the gain-of-function allele, *Sxl^{M#1}*, in which case they enter oogenesis. Conversely, XX germ cells homozygous for a loss-of-function *Sxl* allele enter spermatogenesis even in an ovary, as do XY germ cells (Steinmann-Zwicky et al., 1989). Thus it seems that spermatogenesis by XX germ cells in a testis reflects their failure to activate *Sxl*, because of either a repressive influence from the testicular soma or the absence of an inductive influence normally provided by the ovary. An obvious corollary of this argument is that XY germ cells enter spermatogenesis in an ovary because their *Sxl* gene is inactive, but the presence of *Sxl^{M#1}* was found to be insufficient to

direct XY germ cells into the female pathway in a testis or an ovary. As the authors note, however, this may simply reflect the fact that *Sxl^{M#1}* is not a truly constitutive allele and may be sensitive to upstream signals (see earlier).

A picture that accommodates the available data is that germ line sex determination, like somatic sex determination, depends on the state of activity of *Sxl*. Although somatic *Sxl* activity is determined in a cell-autonomous fashion, the gene's state of activity in the germ line depends on the somatic environment. According to this hypothesis, the pseudomale soma must be less efficient than the male soma at preventing *Sxl* expression, and thus accounts for the observed oogenic stages in these flies.

Although germ line sex determination may depend solely on the state of *Sxl*, complete gametogenesis evidently requires a somatic environment of matching sex, since spermatogenesis in an ovary or oogenesis in a testis is always abortive. However, this does not explain the failure of XX or even XXY germ cells to complete spermatogenesis in a testis, which may rather be due to lack of dosage compensation (Steinmann-Zwicky et al., 1989).

If, as seems likely, other regulatory genes are involved in translating the state of *Sxl* into germ line differentiation, it will be interesting to determine whether they form a hierarchy involving sex-specific splicing decisions.

SUMMARY — A MODEL OF SOMATIC SEX DETERMINATION

Females

- (1) X:A ratio of 1, together with several maternal and zygotic factors, acts to turn on *Sxl*;
- (2) *Sxl* product maintains *Sxl* expression (autoregulatory) and activates *tra*;
- (3) *tra* product, together with a *tra-2* product, induces female-specific expression pattern of *dsx*;
- (4) *dsx* (female) product, together with *ix* product, represses male-specific downstream genes resulting in female sexual phenotype.

Males

- (1) An X:A ratio of 0.5 means that *Sxl* is not activated;
- (2) *tra* is not activated;
- (3) *dsx* is expressed in basal mode;
- (4) *dsx* (male) product represses female-

specific downstream genes resulting in male sexual phenotype.

Alternative states of activity of *Sxl*, *tra* and *dsx* are achieved by sex-specific RNA processing. The function of *tra-2* in somatic sex determination is probably not regulated by alternative splicing, and there are no data for *ix*. Male transcripts are probably nonfunctional in the cases of *Sxl* and *tra*. Male and female transcripts are functional in the case of *dsx*. This summary is shown diagrammatically in Fig. 2.

FUTURE DEVELOPMENTS

To gain further understanding of the sex determination process, it is necessary to clone as many as possible of the terminal differentiation genes that are regulated by the hierarchy, as well as to continue the search for additional regulatory genes. Although the simplest method may at first seem to be differential screening of libraries with male- and female-specific probes, there are many problems inherent in this approach. Clearly, detection in this way of the genes in the hierarchy with alternative sex-specific splicing would not now be predicted, and indeed they were never isolated in screens. But there are also problems with genes that respond to the hierarchy. Genes expressed sex-specifically in sex-specific tissues — namely the gonads — do not seem to be directly under the control of the hierarchy: rather the sex-determining genes are responsible for the decision to develop into male or female gonads, and from then on the genes expressed seem to be programmed by tissue-specific factors.

This class constitutes the vast majority of sex-specific gene products on the evidence of the results of previous screens. The yolk proteins are the only sex-specific genes expressed in a nonsex-specific tissue that have been isolated to date.

So why are we missing these genes? Many genes regulated by the hierarchy may be expressed only transiently, during metamorphosis, at the time when many male or female characteristics are differentiated. This would suggest that screens during this stage may be more fruitful. Furthermore, many genes may not be expressed in a sex-limited way, but may have different spatial or temporal expression patterns, or both, in the two sexes. For example, this would be true of the genes involved in

the differentiation of the male and female abdominal pigmentation patterns or the development of the sex combs.

Can we overcome these problems?

One possibility may be in the recently developed techniques for finding genes expressed in particular patterns by inserting a bacterial β -galactosidase gene near an enhancer for such a gene, and observing the pattern of β -galactosidase staining, which should coincide with the expression of the gene near which it has inserted (O'Kane and Gehring, 1987). Using a series of flies generated by J. Merriam, we have identified a large number of testis-specific patterns of gene expression, some ovary-specific and some specific for other adult cell types. These genes can be cloned readily by virtue of the β -galactosidase gene inserted nearby.

NOTE ADDED IN PROOF

For a review of recent developments in the

story of *daughterless*, see Cline, T. W. 1989. The affairs of *daughterless* and the promiscuity of developmental regulations. *Cell*, 59:231-234.

A recent paper supporting negative regulation of *tra* by *Sxl* is that by Inoue, K., K. Hoshijima, H. Sakamoto, and Y. Shimura. 1990. Binding of the *Drosophila sex-lethal* gene product to the alternative splice site of *transformer* primary transcripts. *Nature*, 344:461-463.

ACKNOWLEDGMENTS

We are grateful to members of our laboratory for discussions and to Janet Panther and Jeanette Ferguson for typing. Several colleagues sent preprints and were extremely helpful and encouraging, particularly Mariana Wolfner, Rolf Nöthiger, Monica Steinmann-Zwicky, Bruce Baker, and Bruce Chase. We are grateful to Simon Holden, Mariana Wolfner, Rolf Nöthiger, Bruce Baker, Bruce Chase, and Pieter Wensink for comments on drafts of this manuscript.

REFERENCES

- Amara, G. G., V. Jonas, M. G. Rosenfeld, E. S. Ong, and R. M. Evans. 1982. Alternative RNA processing in calcitonin gene expression generates mRNAs encoding different polypeptide products. *Nature*, 298: 240-244.
- Amrein, H., M. Gorman, and R. Nöthiger. 1988. The sex-determining gene *tra-2* of *Drosophila* encodes a putative RNA binding protein. *Cell*, 55: 1025-1035.
- Baker, B., and J. Belote. 1983. Sex determination and dosage compensation in *Drosophila melanogaster*. *Annu. Rev. Gen.*, 17: 345-375.
- Baker, B., and K. Ridge. 1980. Sex and the single cell: on the action of major loci affecting sex determination in *Drosophila melanogaster*. *Genetics*, 94: 383-423.
- Baker, B., and M. Wolfner. 1988. A molecular analysis of *doublesex*, a bifunctional gene that controls both male and female sexual differentiation in *Drosophila melanogaster*. *Genes Dev.*, 2: 477-489.
- Barnett, T., C. Pachel, J. P. Gergen, and P. C. Wensink. 1980. The isolation and characterization of *Drosophila* yolk protein genes. *Cell*, 21: 729-738.
- Bell, A. E. 1954. A gene in *Drosophila melanogaster* that produces all male progeny. *Genetics*, 29: 958-959.
- Bell, L. R., E. M. Maine, P. Schedl, and T. W. Cline. 1988. *Sex-lethal*, a *Drosophila* sex-determination switch gene, exhibits sex-specific RNA splicing and sequence similarity to RNA binding proteins. *Cell*, 55: 1037-1046.
- Belote, J. M. 1983. Male-specific lethal mutations of *Drosophila melanogaster*. II. Parameters of gene action during male development. *Genetics*, 105: 881-896.
- Belote, J. M., and B. Baker. 1982. Sex determination in *Drosophila melanogaster*: analysis of *transformer-2*, a sex-transforming locus. *Proc. Natl. Acad. Sci. USA*, 79: 1568-1572.
- , and —. 1983. The dual functions of a sex-determination gene in *Drosophila melanogaster*. *Dev. Biol.*, 95: 512-517.
- Belote, J. M., A. M. Handler, M. F. Wolfner, K. Livak, and B. Baker. 1985. Sex specific regulation of yolk protein gene expression in *Drosophila*. *Cell*, 40: 339-348.
- Belote, J. M., and J. C. Lucchesi. 1980. Male-specific lethal mutations of *Drosophila melanogaster*. *Genetics*, 96: 165-186.
- Belote, J. M., M. B. McKeown, D. J. Andrew, T. N. Scott, M. F. Wolfner, and B. S. Baker. 1985. Control of sexual differentiation in *Drosophila melanogaster*. *Cold Spring Harbor Symp. Quant. Biol.*, 50: 605-614.
- Berleth, T., M. Burri, G. Thoma, D. Bopp, S. Richstein, G. Frigerio, M. Noll, and C. Nüsslein-Volhard. 1988. The role of localisation of *bicoid* RNA in organising the anterior pattern of the *Drosophila* embryo. *EMBO J.*, 7: 1749-1756.
- Bingham, P. M., R. Levis, and G. M. Rubin. 1981. Cloning of DNA sequences from the *white* locus of *D. melanogaster* by a novel and general method.

- Cell*, 25: 693-704.
- Bishop, J. M. 1985. Viral oncogenes. *Cell*, 42: 23-38.
- Boggs, R., P. Gregor, S. Iddriss, J. M. Belote, and M. McKeown. 1987. Regulation of sexual differentiation in *D. melanogaster* via alternative splicing of RNA from the transformer gene. *Cell*, 50: 739-747.
- Bownes, M. 1986. Expression of the genes coding for vitellogenin (yolk protein). *Annu. Rev. Entomol.*, 31: 507-531.
- Bownes, M., and M. Blair. 1986. The effects of a sugar diet and hormones on the expression of the *Drosophila* yolk protein genes. *J. Insect Physiol.*, 32: 493-501.
- Bownes, M., M. Blair, R. Kozma, and M. Dempster. 1983. 20-hydroxyecdysone stimulates tissue-specific yolk-protein gene transcription in both male and female *Drosophila*. *J. Embryol. Exp. Morphol.*, 78: 249-268.
- Bownes, M., A. Scott, and M. Blair. 1987. The use of an inhibitor of protein synthesis to investigate the roles of ecdysteroids and sex-determination genes on the expression of the genes encoding the *Drosophila* yolk proteins. *Development*, 101: 931-941.
- Bownes, M., A. Scott, and A. Shirras. 1988. Dietary components modulate yolk protein gene transcription in *Drosophila melanogaster*. *Development*, 103: 119-128.
- Breathnach, R., and T. Chambon. 1981. Eukaryotic split gene. *Annu. Rev. Biochem.*, 50: 349-383.
- Bridges, C. B. 1921. Triploid intersexes in *Drosophila melanogaster*. *Science*, 54: 252-254.
- . 1925a. Haploidy in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*, 11: 706-710.
- . 1925b. Sex in relation to chromosomes and genes. *Am. Nat.*, 59: 127-137.
- Britten, R., and E. Davidson. 1969. Gene regulation for higher cells: A theory. *Science*, 165: 349-357.
- Brown, E. H., and R. King. 1961. Studies on the expression of the transformer gene in *Drosophila melanogaster*. *Genetics*, 46: 143-156.
- Burtis, K. C., and B. S. Baker. 1989. *Drosophila double-sex* gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. *Cell*, 56: 997-1010.
- Butler, B., V. Pirotta, I. Irminger-Finger, and R. Nöthiger. 1986. The sex-determining gene *tra* of *Drosophila melanogaster*: Molecular cloning and transformation studies. *EMBO J.*, 5: 3607-3613.
- Campuzano, S., L. Carramolina, C. V. Cabrera, M. Ruiz-Gomez, R. Villares, A. Boronat, and J. Modellell. 1985. Molecular genetics of the *achaete-scute* gene complex of *D. melanogaster*. *Cell*, 40: 327-338.
- Caudy, M., H. Vaissin, M. Brand, R. Tuma, L. Y. Jan, and Y. N. Yan. 1988. *daughterless*, a *Drosophila* gene essential for both neurogenesis and sex-determination, has sequence similarities to *myc* and the *achaete-scute* complex. *Cell*, 55: 1061-1067.
- Chandra, H. S. 1985. Sex-determination: a hypothesis based on non-coding DNA. *Proc. Natl. Acad. Sci. USA*, 82: 1165-1169.
- Chapman, K. B., and M. F. Wolfner. 1988. Determination of male specific gene expression in *Drosophila* accessory glands. *Dev. Biol.*, 126: 195-202.
- Cline, T. W. 1976. A sex-specific, temperature-sensitive maternal effect of the *daughterless* mutation of *Drosophila melanogaster*. *Genetics*, 84: 723-742.
- . 1978. Two closely linked mutations in *Drosophila melanogaster* that are lethal to opposite sexes and interact with *daughterless*. *Genetics*, 90: 683-698.
- . 1979a. A product of the maternally-influenced *Sex-lethal* gene determines sex in *Drosophila melanogaster*. *Genetics*, 91: 22.
- . 1979b. A male-specific lethal mutation in *Drosophila melanogaster* that transforms sex. *Dev. Biol.*, 72: 266-275.
- . 1980. Maternal and zygotic sex-specific gene interactions in *Drosophila melanogaster*. *Genetics*, 96: 903-926.
- . 1983. The interaction between *daughterless* and *Sex-lethal* in triploids: A lethal sex-transforming maternal effect linking sex-determination and dosage compensation in *Drosophila melanogaster*. *Dev. Biol.*, 95: 260-274.
- . 1984. Autoregulatory functioning of a *Drosophila* gene product that establishes and maintains the sexually determined state. *Genetics*, 107: 231-277.
- . 1985. Primary events in the determination of sex in *Drosophila melanogaster*. In H. O. Halvorson and A. Monroy (eds.), *Origin and Evolution of Sex*, pp. 301-327. Alan R. Liss, New York.
- . 1986. A female-specific lethal lesion in an X-linked positive regulator of the *Drosophila* sex determination gene, *Sex-lethal*. *Genetics*, 113: 641-663.
- . 1988. Evidence that *sisterless-a* and *sisterless-b* are two of several discrete "numerator elements" of the X/A sex determination signal in *Drosophila* that switch *Sxl* between two alternative stable expression states. *Genetics*, 119: 829-862.
- Cronmiller, C., and T. W. Cline. 1987. The *Drosophila* sex determination gene *daughterless* has different functions in the the germ line vs. the soma. *Cell*, 48: 479-487.
- Cronmiller, C., P. Schedl, and T. W. Cline. 1988. Molecular characterisation of *daughterless*, a *Drosophila* sex-determination gene with multiple roles in development. *Genes Dev.*, 2: 1666-1676.
- Davis, R. L., H. Weintraub, and A. B. Lassar. 1987.

- Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell*, 51: 987-1000.
- Di Benedetto, A. J., D. M. Lakich, W. D. Kruger, J. M. Belote, B. S. Baker, and M. F. Wolfner. 1987. Sequences expressed sex-specifically in *Drosophila melanogaster* adults. *Dev. Biol.*, 719: 242-251.
- Ferguson, E. L., P. W. Sternberg, and H. R. Horvitz. 1987. A genetic pathway for the specification of vulval cell lineages of *Caenorhabditis elegans*. *Nature*, 326: 259-267.
- Frigerio, G., M. Burri, D. Bopp, S. Baumgartner, and M. Noll. 1986. Structure of the segmentation gene *paired* and the *Drosophila* PRD gene set as part of a gene network. *Cell*, 47: 735-746.
- Gadagkar, R., V. Nanjundiah, N. Joshi, and H. S. Chandra. 1982. Dosage compensation and sex-determination in *Drosophila*. Mechanism of measurement of the X:A ratio. *J. Biosci.*, 4: 377-390.
- Gans, M., C. Audit, and M. Masson. 1975. Isolation and characterisation of sex-linked female sterile mutants in *Drosophila melanogaster*. *Genetics*, 81: 683-704.
- Garabedian, M. J., M. C. Hung, and P. C. Wensink. 1985. Independent control elements that determine yolk protein gene expression in alternative *Drosophila* tissues. *Proc. Natl. Acad. Sci. USA*, 82: 1396-1400.
- Garabedian, M. J., B. M. Shepherd, and P. C. Wensink. 1986. A tissue specific transcription enhancer from the *Drosophila* Yolk Protein 1 gene. *Cell*, 45: 859-867.
- Garabedian, M. J., A. D. Shirras, M. Bownes, P. C. Wensink. 1987. The nucleotide sequence of the gene coding for *Drosophila melanogaster* yolk protein 3. *Gene*, 55: 1-8.
- Garcia-Bellido, A. 1979. Genetic analysis of the *achaete-scute* system of *Drosophila melanogaster*. *Genetics*, 91: 491-520.
- Gehring, W. J. 1976. Developmental genetics of *Drosophila*. *Annu. Rev. Genet.*, 10: 209-252.
- Gehring, W. J., R. Klemenz, U. Weber, and U. Kloter. 1984. Functional analysis of the *white* gene of *Drosophila* by P-factor mediated transformation. *EMBO J.*, 3(9): 2077-2085.
- Gergen, J. P. 1987. Dosage compensation in *Drosophila*: evidence that *daughterless* and *sex-lethal* control X chromosome activity at the blastoderm stage of embryogenesis. *Genetics*, 117: 477-485.
- Gollin, S. M., and R. C. King. 1981. Studies for *fs(1)1621*, a mutation producing ovarian tumours in *Drosophila melanogaster*. *Dev. Genet.*, 2: 203-218.
- Goralski, T. J., J. E. Edström, and B. S. Baker. 1989. The sex-determination locus *transformer-2* of *Drosophila* encodes a polypeptide with similarity to RNA binding proteins. *Cell*, 56: 1011-1018.
- Hannah-Alava, A., and C. Stern. 1957. The sexcomb in males and intersexes of *Drosophila melanogaster*. *J. Exp. Zool.*, 134: 533-555.
- Hildreth, P. E. 1965. *Doublesex*, a recessive gene that transforms both males and females of *Drosophila* into intersexes. *Genetics*, 51: 659-678.
- Hung, M. C., T. Barnett, C. Woolford, and P. C. Wensink. 1982. Transcript maps of the *Drosophila* yolk protein genes. *J. Mol. Biol.*, 154: 581-602.
- Jowett, T., and J. Postlethwait. 1980. The regulation of yolk protein polypeptide synthesis in *Drosophila* ovaries and fat body by 20-hydroxyecdysone and a juvenile hormone analogue. *Dev. Biol.*, 80: 225-234.
- Kaufman, T. C., R. A. Lewis, and B. C. Wakimoto. 1980. Cytogenetic analysis of Chromosome 3 in *Drosophila melanogaster*. The homeotic gene complex in polytene chromosome interval 84A-B. *Genetics*, 94: 115-120.
- Kingston, R. E., A. S. Baldwin, Jr., and P. A. Sharp. 1984. Regulation of heat shock protein 70 gene expression by *c-myc*. *Nature*, 312: 280-282.
- Krumm, A., L. E. Roth, and G. Korge. 1985. Transformation of salivary gland secretion protein gene *Sgs-4* in *Drosophila*: Stage and tissue-specific regulation, dosage compensation and position effect. *Proc. Natl. Acad. Sci. USA*, 82: 5055-5059.
- Lehmann, R., and C. Nüsslein-Volhard. 1986. Abdominal segmentation, pole cell formation and embryonic polarity requires the localised activity of *oskar*, a maternal gene in *Drosophila*. *Cell*, 47: 141-152.
- Lewis, E. B. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature*, 276: 565-570.
- Lindsley, D. L., and E. H. Grell. 1968. Genetic variations of *Drosophila melanogaster*. *Carnegie Inst. Wash. Publ.*, 627.
- Lucchesi, J. C., and T. Skripsky. 1981. The link between dosage compensation and sex differentiation in *Drosophila melanogaster*. *Chromosoma*, 82: 217-227.
- Maine, E. M., H. K. Salz, P. Schedl, and T. W. Cline. 1986. *Sex-lethal*, a link between sex-determination and sexual differentiation in *Drosophila melanogaster*. *Cold Spring Harbor Symp. Quant. Biol.*, 50: 595-604.
- Mariani, B. D., J. R. Lingappa, and F. C. Kafatos. 1988. Temporal regulation in development: negative and positive *cis* regulators dictate the precise timing of expression of a *Drosophila* chorion gene. *Proc. Natl. Acad. Sci. USA*, 85: 3029-3033.
- Marsh, J. L., and E. Wieschaus. 1978. Is sex determination in germline and soma controlled by separate genetic mechanisms? *Nature*, 272: 249-251.
- Marshall, T., and J. R. Whittle. 1978. Genetic analysis of the mutation female-lethal in *Drosophila melanogaster*. *Genet. Res.*, 32: 103-111.
- McKeown, M., J. M. Belote, and B. S. Baker. 1987. A molecular analysis of *transformer*, a gene in *Drosophila melanogaster* that controls female sexual differentiation. *Cell*, 48: 489-499.

- McKeown, M., J. M. Belote, and R. T. Boggs. 1988. Ectopic expression of the *transformer* gene products leads to female differentiation of chromosomally male *Drosophila*. *Cell*, 53: 887-895.
- Monsma, S. A., and M. F. Wolfner. 1988. Structure and expression of a *Drosophila* male accessory gland gene whose product resembles a peptide pheromone precursor. *Genes Dev.*, 2: 1063-1073.
- Morgan, T. H., H. Redfield, and L. V. Morgan. 1943. Maintenance of a *Drosophila* stock centre in connection with investigations on the constitution of germinal material in relation to heredity. *Carnegie Inst. Wash. Yearbook*, 42: 171-174.
- Muller, H. J., and S. Zimmering. 1960. A sex-linked lethal without evident effect in *Drosophila* males, but partially dominant in females. *Genetics*, 45: 1000-1002.
- Murre, C., P. S. McCaw, and D. Baltimore. 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, *daughterless*, *MyoD* and *myc* proteins. *Cell*, 56: 777-783.
- Nagoshi, R., M. McKeown, K. Burtis, J. M. Belote, and B. S. Baker. 1988. The control of alternative splicing at genes regulating sexual differentiation in *Drosophila melanogaster*. *Cell*, 53: 229-236.
- Nöthiger, R., M. Jonglez, M. Leuthold, P. Meier-Gerschwiler, and T. Weber. 1989. Sex-determination in the germ line of *Drosophila* depends on genetic signals and inductive somatic factors. *Development*, 107: 505-518.
- Nöthiger, R., M. Leuthold, P. Meier-Gerschwiler, A. Grünen, W. Keller, C. Leist, M. Roost, and H. Schmid. 1987. Genetic and developmental analysis of the sex-determining gene *doublesex* (*dsx*) of *Drosophila melanogaster*. *Genet. Res.*, 50: 113-123.
- Nöthiger, R., and M. Steinmann-Zwicky. 1985. Sex determination in *Drosophila*. *Trends Genet.*, 1: 209-214.
- , and ———. 1987. Genetics of sex-determination: what can we learn from *Drosophila*? *Development* (Suppl.), 101: 17-24.
- Oliver, B., N. Perrimon, and A. Mahowald. 1988. Genetic evidence that the *sans fille* locus is involved in *Drosophila* sex-determination. *Genetics*, 120: 159-171.
- O'Kane, C. J., and W. J. Gehring. 1987. Detection *in situ* of genomic regulatory elements in *Drosophila*. *Proc. Natl. Acad. Sci. USA*, 84: 9123-9127.
- Perrimon, N., J. D. Mohler, L. Engstrom, and A. P. Mahowald. 1986. X linked female sterile loci in *Drosophila melanogaster*. *Genetics*, 113: 695-712.
- Pirrotta, V., H. Steller, and M. P. Bozzetti. 1985. Multiple upstream regulatory elements control the expression of the *Drosophila white* gene. *EMBO J.*, 4: 3501-3508.
- Postlethwait, J. H., M. Bownes, and T. Jowett. 1980. Sexual phenotype and vitellogenin synthesis in *Drosophila melanogaster*. *Dev. Biol.*, 79: 379-387.
- Ptashne, M. 1986. Gene regulation by proteins acting nearby and at a distance. *Nature*, 322: 697.
- Romano, C. P., B. Bienz-Tadmor, B. D. Mariani, and F. C. Kafatos. 1988. Both early and late *Drosophila* chorion gene promoters confer correct temporal, tissue and sex-specificity on a reporter *Adh* gene. *EMBO J.*, 7: 783-790.
- Rubin, G. M., and A. C. Spradling. 1982. Genetic transformation of *Drosophila* with transposable element vectors. *Science*, 218: 348-353.
- Salz, H. K., T. W. Cline, and P. Schedl. 1987. Functional changes associated with structural alterations induced by mobilisation of a P-element inserted in the *Sex-lethal* gene of *Drosophila*. *Genetics*, 117: 221-231.
- Salz, H. K., E. M. Maine, L. N. Keyes, M. E. Samuels, T. W. Cline, and P. Schedl. 1989. The *Drosophila* female specific sex-determination gene, *Sex-lethal*, has stage-, tissue-, and sex-specific RNAs suggesting multiple modes of regulation. *Genes Dev.*, 3: 708-719.
- Sanchez, L., and R. Nöthiger. 1982. Clonal analysis of *Sxl*, a gene needed for female sexual development in *Drosophila melanogaster*. *Wilhelm Roux's Arch. Dev. Biol.*, 191: 211-214.
- , and ———. 1983. Sex determination and dosage compensation in *Drosophila melanogaster*: Production of male clones in XX females. *EMBO J.*, 2: 485-491.
- Schäfer, U. 1986. The regulation of male specific transcripts by sex-determination genes in *Drosophila melanogaster*. *EMBO J.*, 5: 3579-3583.
- Schleif, R. 1987. Why should DNA loop? *Nature*, 327: 369-370.
- Schüpbach, T. 1982. Autosomal mutations that interfere with sex-determination in somatic cells of *Drosophila melanogaster* have no direct effect on the germline. *Dev. Biol.*, 89: 117-127.
- . 1985. Normal female germ cell differentiation requires the female X chromosome to autosome ratio and expression of *Sex-lethal* in *Drosophila melanogaster*. *Genetics*, 109: 529-548.
- Seidel, S. 1963. Experimentelle Untersuchungen über die Grundlagen der Sterilität von *transformer* (*tra*) Männchen bei *Drosophila melanogaster*. *Z. Vererbungsl.*, 94: 215-241.
- Shirras, A. D., and M. Bownes. 1987. Separate DNA sequences are required for normal female and ecdysone-induced male expression of *Drosophila melanogaster* yolk protein 1. *Mol. Gen. Genet.*, 210: 153-155.
- , and ———. 1989. *cricklet*: a locus regulating a number of adult functions of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*, 86: 4559-4563.
- Skripsky, T., and J. C. Lucchesi. 1982. Intersexuality resulting from the interaction of sex-specific

- lethal mutations in *Drosophila melanogaster*. *Dev. Biol.*, 94: 153-162.
- Sosnowski, B. A., J. M. Belote, and M. McKeown. 1989. Sex-specific alternative splicing of RNA from the *transformer* gene results from sequence-dependent splice site blockage. *Cell*, 58: 449-459.
- Spradling, A. C., and A. P. Mahowald. 1980. Amplification of genes for chorion proteins during oogenesis in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*, 77: 1096-1100.
- Steinmann-Zwicky, M. 1988. Sex-determination in *Drosophila*: the X chromosomal gene *liz* is required for *Sxl* activity. *EMBO J.*, 7: 3889-3898.
- Steinmann-Zwicky, M., and R. Nöthiger. 1985. A small region on the X chromosome of *Drosophila* regulates a key gene that controls sex-determination and dosage compensation. *Cell*, 42: 877-887.
- Steinmann-Zwicky, M., H. Schmid, and R. Nöthiger. 1989. Cell autonomous and inductive signals can determine the sex of the germ line of *Drosophila* by regulating the gene *Sxl*. *Cell*, 57: 157-166.
- Stern, C. 1966. Pigmentation mosaicism in intersexes of *Drosophila*. *Rev. Suisse Zool.*, 73: 339-355.
- Sturtevant, A. H. 1945. A gene in *Drosophila melanogaster* that transforms females into males. *Genetics*, 30: 297-299.
- Sullivan, N. F., R. A. Watt, M. R. Delaney, C. L. Green, and D. L. Spector. 1986. Colocalisation of the *myc* oncogene protein and small nuclear ribonucleoprotein particles. *Cold Spring Harbor Symp. Quant. Biol.*, 51: 943-947.
- Sutcliffe, J. G., and R. J. Milner. 1988. Alternative mRNA splicing: the *Shaker* gene. *Trends Genet.*, 4: 297-299.
- Tanaka, Z., A. Fukunaga, and K. Oishi. 1976. Studies on the sex-specific lethals of *Drosophila melanogaster*: further studies on male-specific lethal gene, *maleless*. *Genetics*, 84: 257-266.
- Tompkins, L. 1986. Genetic control of behaviour in *Drosophila melanogaster*. *Trends Genet.*, 2: 14-17.
- Torres, M., and L. Sanchez. 1989. The *scute* (T4) gene acts as a numerator element of the X:A signal that determines the state of activity of *Sex-lethal* in *Drosophila*. *EMBO J.*, 8: 3079-3086.
- Uchida, S., T. Uenoyama, and K. Oishi. 1981. Studies on the sex-specific lethals of *Drosophila melanogaster*. III. A third chromosome male-specific lethal mutant. *Jpn. J. Genet.*, 56: 523-527.
- Uenoyama, T., A. Fukunaga, and K. Oishi. 1982. Studies on the sex-specific lethals of *Drosophila melanogaster*. IV. Gynandromorph analysis of three male-specific lethals, *mle*, *msl-2²⁷* and *mle(3)I32*. *Genetics*, 102: 223-231.
- Van Deusen, E. B. 1976. Sex-determination in germ-line chimeras of *Drosophila melanogaster*. *J. Embryol. Exp. Morphol.*, 37: 173-185.
- Van Handel, E., and A. D. Lea. 1984. Vitellogenin synthesis in blood fed *Aedes aegypti* in the absence of the head, thorax and ovaries. *J. Insect Physiol.*, 30: 871-875.
- Watanabe, T. K. 1975. A new sex-transforming gene on the second chromosome of *Drosophila melanogaster*. *Jpn. J. Genet.*, 50: 269-271.
- Wolfner, M. F. 1988. Sex-specific gene expression in somatic tissues of *Drosophila melanogaster*. *Trends Genet.*, 4: 333-337.
- Wong, Y. C., J. Pustell, N. Spoerel, and F. C. Kafatos. 1985. Coding and potential regulatory sequences of a cluster of chorion genes in *Drosophila melanogaster*. *Chromosoma*, 92: 124-135.